An EnKF-based Flow State Estimator for Aerodynamic Problems

Thesis by
Andre F. C. da Silva

In Partial Fulfillment of the Requirements for the
degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2019
Defended September 5th, 2018

To my lovely wife.

Without her help, sacrifice and encouragement,

it simply would've never been.

"It is the glory of God to conceal things,

but the glory of kings is to search things out."

Prov. 25:2 (ESV)

ACKNOWLEDGEMENTS

I'm also thankful for the care and hospitality of the brothers and sisters from Calvary OPC and Pasadena OPC. Their display of the Christian love truly made the solitude of the last nine months bearable.

Finally, but not least, I'm forever indebted to my extended family and friends back in Brazil and elsewhere. Those are the ones that cheered, prayed and willed me through this doctorate. Every small token of remembrance and affection were of extreme importance for me.

ABSTRACT

Regardless of the plant model, robust flow estimation based on limited measurements remains a major challenge in successful flow control applications. Aiming to combine the robustness of a high-dimensional representation of the dynamics with the cost efficiency of a low-order approximation of the state covariance matrix, a flow state estimator based on the Ensemble Kalman Filter (EnKF) is applied to two-dimensional flow past a cylinder and an airfoil at high angle of attack and low Reynolds number. For development purposes, we use the numerical algorithm as both the estimator and as a surrogate for the measurements. In a perfect-model framework, a reduced number of either pressure sensors on the surface of the body or sparsely placed velocity probes in the wake are sufficient to accurately estimate the instantaneous flow state. Because the dynamics of these flows are restricted to a low-dimensional manifold of the state space, a small ensemble size is sufficient to yield the correct asymptotic behavior. The relative importance of each sensor location is evaluated by analyzing how they influence the estimated flow field, and optimal locations for pressure sensors are determined.

However, model inaccuracies are ubiquitous in practical applications. Covariance inflation is used to enhance the estimator performance in the presence of unmodeled freestream perturbations. A combination of parametric modeling and augmented state methodology is used to successfully estimate the forces on immersed bodies subjected to deterministic and random gusts. The robustness of high-dimensional representation of the dynamics to the choice of parameters such as the Reynolds number is inherited by the estimator, which was shown to successfully estimate the reference Reynolds number on the fly. Spatial and temporal discretization can constitute a second source of errors which can render numerical solutions a biased representation of reality. Left unaccounted for, biased forecast and observation models can lead to poor estimator performance. In this work, we propose a low-rank representation for the bias whose dynamics are represented by a colored-noise process. System state and bias parameters are simultaneously tracked online with the Ensemble Kalman Filter (EnKF) algorithm. The proposed methodology is demonstrated to achieve a 70% error reduction for the problem of estimating the state of the two-dimensional low-Re flow past a flat plate at high angle of attack using an ensemble of coarse-mesh simulations and pressure measurements at the surface of the body, compared to a bias-blind estimator. Strategies to determine

the bias statistics and to deal with nonlinear observation functions in the context of ensemble methods are discussed.

PUBLISHED CONTENT AND CONTRIBUTIONS

[1] D. Darakananda et al. "EnKF-based Dynamic Estimation of Separated Flows with a Low-Order Vortex Model". In: *2018 AIAA Aerospace Sciences Meeting*. AIAA 2018-0811. Reston, Virginia: American Institute of Aeronautics and Astronautics, Jan. 2018. DOI: 10.2514/6.2018-0811.
A.F.C.S. participated in the conception and implementation of the estimator, and participated in interpretation of the results.

[2] A. F. C. da Silva and T. Colonius. "A Bias-aware EnKF Estimator for Aerodynamic Flows". In: *48th AIAA Fluid Dynamics Conference*. AIAA 2018-3225. Reston, Virginia: American Institute of Aeronautics and Astronautics, June 2018. DOI: 10.2514/6.2018-3225.
A.F.C.S. participated in the conception and implementation of the estimator, obtained the simulation data, and was the principle author of the manuscript.

[3] A. F. C. da Silva and T. Colonius. "Ensemble-Based State Estimator for Aerodynamic Flows". In: *AIAA Journal* (June 2018), pp. 1–11. DOI: 10.2514/1.J056743.
A.F.C.S. participated in the conception and implementation of the estimator, obtained the simulation data, and was the principle author of the manuscript.

[4] X. An et al. "Response of the Separated Flow over an Airfoil to a Short-Time Actuator Burst". In: *47th AIAA Fluid Dynamics Conference*. AIAA 2017-3315. Reston, Virginia: American Institute of Aeronautics and Astronautics, June 2017. DOI: 10.2514/6.2017-3315.
A.F.C.S. participated in the conception of the project, and participated in interpretation of the results.

[5] A. F. C. da Silva and T. Colonius. "An EnKF-based Flow State Estimator for Aerodynamic Flows". In: *8th AIAA Theoretical Fluid Mechanics Conference*. AIAA 2017-3483. Reston, Virginia: American Institute of Aeronautics and Astronautics, June 2017. DOI: 10.2514/6.2017-3483.
A.F.C.S. participated in the conception and implementation of the estimator, obtained the simulation data, and was the principle author of the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

NOMENCLATURE

α.   multiplicative covariance inflation parameter.

β.   additive covariance inflation parameter.

Γ_x, Γ_y.   respectively, dynamics and observation bias low-rank models.

$\mathbb{1}$.   1-dimensional vector of ones.

$\mathbb{E}[\cdot]$.   expectation operator defined in Eq. 1.3.

μ_k, ν_k.   process (i.i.d. $N(0, Q)$) and observation (i.i.d. $N(0, R)$) noises.

$\rho(\cdot)$.   probability density function.

ξ_k, η_k.   respectively, dynamics and observation bias parameters.

A_k.   state perturbation matrix (see Eq. 5.16) at time t_k.

c.   dimension of the control vector.

C_k^{xy}.   cross-covariance matrix between variables x and y at time t_k.

E_x, E_y.   state and observation estimate errors, respectively.

$f : \mathbb{R}^n \mapsto \mathbb{R}^n$.   (possibly nonlinear) dynamic model.

$h : \mathbb{R}^n \mapsto \mathbb{R}^p$.   (possibly nonlinear) observation model.

$M(\tau)$.   autocorrelation matrix with lag τ.

n.   dimension of the system state vector.

n_o.   dimension of the observation bias parameters vector.

n_s.   dimension of the state bias parameters vector.

p.   dimension of the measurement/observation vector.

q.   ensemble size.

Re.   Reynolds number $\frac{U_\infty L}{\nu}$, where L is a reference length.

RMS_x, RMS_y.   state and observation ensemble root mean square errors, respectively.

v_k.   analysis correction coefficients.

x_k.   system state vector at time t_k.

y_k.   vector of measured quantities at time t_k.

Z_k.   state ensemble matrix (each column represents the state of the corresponding ensemble member).

Chapter 1

INTRODUCTION

1.1 Motivation

In the aeronautical context, unsteady conditions, such as the ones that would result from a maneuver or from the occurrence of gusts, are ever present. The agility and maneuverability of a fighter or the perceived level of comfort of a commercial aircraft could be significantly enhanced if we could robustly estimate the instantaneous flow state from the available measurement data (e.g. pressure readings on the wings and fuselage), and act accordingly using closed-loop flow control[1]. In this work we focus on one of the two key ingredients to any successful closed-loop control design: the ability to predict the state of a fluid system and forecast its evolution.

Figure 1.1 can be used to describe how different scientific communities have been approaching the dilemma between model complexity (x-axis) and estimation rate (y-axis). Estimation rate refers to the number of forecasts per unit time, and, for on-line estimation, is set by the system's dynamics but highly constrained by the available computational power. The gray area between the axis represents the choices of model complexity and estimation rate that are achievable with the currently available computational power. The horizontal dashed line represents the minimum estimation rate that would allow us real-time prediction. Because standard estimation techniques don't scale well with an increasing number of degrees of freedom, the control engineering community will generally favor low-rank models that preserve limited, but dynamically important, features of the system. On the other hand, fluid mechanicists, particularly those in the CFD community, typically use every addition to the available computational power to simulate models that are more complex and reliable than their predecessors, even if these simulations take months.

If performing closed-loop control is the end goal of this estimation, then the real-time requirement is fundamental, and only the solutions that lie above the horizontal dashed line are ultimately useful. In the flow control community, one common approach is to use dimensionality reduction techniques such as Balanced Truncation[2] or Eigenvalue Realization Algorithm [3] to highlight the most important features of the dynamical system under scrutiny. The resulting reduced-order models (ROM)

can be made small enough to allow the use of the standard algorithms, but their well-known fragility to the specification of initial conditions and flow parameters (e.g., Reynolds number) can constitute a major applicability limitation. Therefore, it would be desirable to seek more robust solutions that lie close to the intersection of the real-time barrier with the computational power barrier.
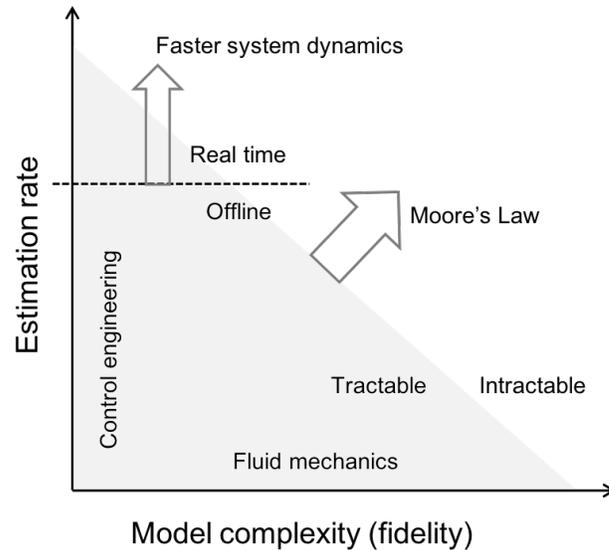
Figure 1.1: Schematics on the current development of estimation techniques in the fluid mechanics context.

Although several state estimation techniques have been developed, their application to fluid dynamics is challenging due to the nonlinearity and high dimensionality of the underlying physical phenomena. Fluid systems are represented by spatially continuous models and any suitable discretization results in high-dimensional discrete models. Whereas feedback control applications require real-time state estimation, the computational cost of standard control techniques such as the Kalman filter [4] do not scale well with increasing flow complexity and faster time scales. Also, numerical simulations require appropriate boundary and initial conditions that are often uncertain[5]. Measurement data can provide the necessary information to close the gap between simulation and experiments. The development of methodologies for the seamless integration of measurement data and (often sophisticated) mathematical models is the goal of a research area know as *data assimilation* (DA).

1.2 Bibliographic Review

Table 1.1 summarizes recent works on data-driven flow estimation. Two main goals have driven the fluid mechanics community to combine numerical models and experimental data: flow reconstruction and flow prediction.

The first application is as a form of extending the experimental reach by performing what is commonly referred to as *enhanced experiment* or *hybrid simulation*. As pointed out by Nisugi, Hayase, and Shirai [6], taken individually, computational fluid dynamics (CFD) and experimental data lack the ability to fully represent the physical phenomena under scrutiny. Despite being a direct observation of the true physics, experimental data is inevitably corrupted by noise and only a small subset of the relevant physical information can be measured[1]. In addition, there can be uncertainties and biases in the transduction process. On the other hand, limited computational resources restrict our ability to accurately represent the underlying flow physics, especially for higher Reynolds numbers. Moreover, assumptions regarding the initial and boundary conditions are often too simplistic to accurately represent the conditions that would be encountered in an experimental setting. But, despite all these limitations, a numerical solution allows for the evaluation of physical quantities that are unattainable using instrumentation. By combining the strengths of both approaches, the resulting hybrid solution is able to provide flow information that is consistent with the observed experimental data and recovers quantities that were not directly observed in the experiment Hayase [7].

Nisugi *et al.* [6] and Hayase *et al.* [5] incorporated measurement data into a simulation using a feedback controller whose constant gain was determined by trial and error to obtain the best fit to experimental data. Other researchers formulated this problem from a variational perspective, in which they seek to minimize a cost functional describing the data mismatch subjected to constraints[2][8]. Papadakis and Memin [9] and Gronskis, Heitz, and Memin [10] treated the system dynamics as deterministic, and used a variational framework to estimate the initial and boundary conditions that were most consistent with the available observations.

Suzuki et al. [11, 12] examined the problem of estimating the flow past a NACA 0012 airfoil for different angles of attack and Re between 10^3 and 10^4. Their

---

[1]This limitation can manifest itself in many ways: limited field of view, limited temporal and spatial resolution, methodological inability to measure some quantities (e.g., vorticity or Reynolds stresses), etc.

[2]The constraints can be dynamical, such as enforcing the estimate to be a solution of the Navier-Stokes equations, or kinematic, such as enforcing the divergence-free condition.

assimilation approach consisted in taking a weighted averaged of the DNS solution and a rectified version of the experimental data. The rectification procedure ensured that the added data satisfied the divergence-free condition. Their conclusion was that the hybrid simulation performed better[3] than an unsteady Reynolds-averaged Navier-Stokes (URANS) simulation for higher angles of attack. Foures et al. [13] and Symon et al. [14] used a variational method to incorporate observations of the mean velocity profile into a RANS simulation in order to reduce the amount of measurement noise and estimate physical quantities, such as the Reynolds shear stress, that were not directly measured.

The second goal of data assimilation has closed-loop control in mind. In this case, the estimator accuracy requirements must be weighed against turnaround time (see Fig. 1.1). Within the flow control community, the most common approach to flow estimation is the development of reduced-order models (ROMs) whose number of degrees of freedom are small enough to be tractable with the classical estimation techniques. Provided these models have a low number of degrees of freedom, then the classical control techniques become feasible. Gerhard et al. [15] used a 3-mode POD-Galerkin model (enhanced with the shift mode) to design a controller to suppress vortex shedding behind circular cylinders at low Reynolds numbers. Aleksic et al. [16] used data from 15 pressure sensors and a 5-mode Galerkin model to decrease and stabilize the drag of a 2D bluff body. Ahuja and Rowley [2] used a 22-mode ROM obtained by Balanced Truncation to design a LQG controller for the flow past an inclined flat plate. Flinois and Morgans [3] used the Eigenvalue Realization Algorithm (ERA) to construct an unstable 10-mode ROM which was then used to design $\mathcal{H}_\infty$ controllers to stabilize the system. These ROMs, however, are fragile with respect to initial conditions and flow parameters like the Reynolds number[1]. Alternatively, researchers such as Fukumori and Malanotte-Rizzoli [17] and Suzuki [18] proposed the use of reduced-order approximations to the Kalman filter that restrict the correction to the larger scales of the solution and alleviate the computational cost involved.

As an alternative approach, researchers in fields such as meteorology, oceanography and geophysical fluid dynamics have developed data assimilation algorithms that are inherently capable of dealing with high-dimensional nonlinear systems and high volumes of measurement data [19, 20]. These methods take advantage of the increasingly available computational power and efficient parallel implementations,

---

[3]The hybrid simulation produced better estimates for the lift coefficient and the dynamics of the vortical structures in the shear layer.

and had not, until recently, received much attention from the flow control community. A three-paper series by Bewley and his collaborators aimed to apply Kalman filtering to devise a state estimator for laminar [21] and turbulent channel flow[22, 23]. Following a rigorous derivation of stochastic models for the system noises, they were able to successfully track the wall-normal velocity and vorticity based on pressure and wall skin friction. For the turbulent channel flows, an Ensemble Kalman Filter was shown to achieve at least one order of magnitude less error than previously reported in the literature at 20 viscous units from the wall. Around the same time, Kato and Obayashi [24] used the Ensemble Kalman Filter to estimate the velocity field behind a square cylinder by assimilation of pressure measurements at the faces of the body. These two papers appear to be the first application of ensemble-based estimation methods to classical fluid dynamics problems.

Recent applications of the EnKF include Kikuchi, Misaka, and Obayashi [25], who compared the performance of a EnKF and a Particle Filter applied to a POD-Galerkin model of the problem of the flow past a cylinder, and Kato et al. [26], who used a variation of the EnKF to achieve synchronization between a RANS numerical simulation of a steady transonic flow past airfoils and pressure experimental data. Mons et al. [27] used an ensemble Kalman smoother and other ensemble-based variational methods to reconstruct freestream perturbation history based on measurements taken on and around a circular cylinder subjected to it.

While the use of ensemble methods make the estimation of complex flow systems viable, it also renders the estimator more vulnerable to the deleterious effects of modeling errors. However, past works in flow estimation did very little in proposing ways of quantify and mitigate these effects, and the advent of a robust and efficient error control methodology could potentially improve the performance of sequential estimators.

Table 1.1: Summary of the recent contributions to the area of flow estimation.

Reference	IP/DA[4]	Methodology	Estimated Flow		Estimated Quantities	Estimator Model	Measurements
			Description	Re[5]			
Bewley and Protas [28]	DA	Adjoint	Incompressible 3D turbulent channel flow	100	Initial State	Linearized N-S	Skin friction and pressure distribution
Hayase, Nisugi, and Shirai [5] and Nisugi, Hayase, and Shirai [6]	DA	Constant-gain observer	Flow past a square cylinder	1200	Velocity field	DNS	Pressure on square faces
Hoepffner et al. [21]	DA	KF/EKF	3D Laminar channel flow	3000	Wall-normal velocity and vorticity	Linearized N-S	Skin friction and pressure distribution
Chevalier et al. [22]	DA	EKF	3D Turbulent channel flow	100	Wall-normal velocity and vorticity	Linearized N-S	Skin friction and pressure distribution
Ruhnau, Stahl, and Schnorr [29]	DA	Adjoint	Synthetic turbulent flow	10	Vorticity field	DNS	Optical flow
Papadakis and Memin [9]	DA	Adjoint	Incompressible 2D turbulent field	Not re- ported	Initial State	DNS	Velocity field

Continues on next page...

Table 1.1 – *Continued from previous page*

| Reference | IP/DA | Methodology | Estimated Flow | | Estimated Quantities | Estimator Model | Measurements |
			Description	Re			
Suzuki, Ji, and Yamamoto [11] and Suzuki et al. [12]	DA	Weighted average	Flow past a NACA 0012	$10^3 - 10^4$	Velocity field	DNS	TR-PTV
Colburn, Cessna, and Bewley [23]	DA	EnKF	3D Turbulent channel flow	100	Wall-normal velocity and vorticity	DNS	Skin friction and pressure distribution
Kato and Obayashi [24]	DA	EnKF	Flow past a square cylinder	1200	Velocity field	DNS	Pressure on square faces
Suzuki [18]	DA	Low-rank Approx. EKF	Planar jet flow	2000	Velocity field	DNS	TR-PTV
Gronskis, Heitz, and Memin [10]	DA	Adjoint	Flow past a circular cylinder	172	Boundary and initial conditions	DNS	Vorticity field
Foures et al. [13]	IP	Adjoint	Flow past a circular cylinder	150	Reynolds stress tensor	RANS	Velocity field
Kato et al. [26]	IP	EnKF	Transonic flow past an airfoil and a wing	$10^6 - 10^7$	Flow variables + turbulent viscosity	RANS	c_p on surfaces

Continues on next page...

Table 1.1 – *Continued from previous page*

Reference	IP/DA	Methodology	Estimated Flow		Re	Estimated Quantities	Estimator Model	Measurements
			Description					
Kikuchi, Misaka, and Obayashi [25]	DA	EnKF/PF	Flow past a circular cylinder		1000	ROM coeffs	POD-Galerkin ROM	Velocity in the wake
Mons et al. [27]	DA	EnKS/4DVar/ 4DEnVar	Circular cylinder subjected to gusts		100	Boundary and initial conditions	DNS	c_p, C_D, C_L and velocity field
Mons, Chassaing, and Sagaut [30]	DA	Adjoint	Rotating cylinder		100	Initial State and Ω	DNS	C_D, C_L, and velocity field
Meldi and Poux [31]	DA	Filtered-Covariance KF	Circular Cylinder/Thick Plate		100/ 80000	Velocity and pressure field	DNS/DDES	Velocity field
Symon et al. [14] and Symon [32]	IP	Adjoint	Flow past an airfoil		13500	Reynolds stress tensor	RANS	Velocity field
Darakananda et al. [33]	DA	EnKF	Inclined flat plate		500	Vortices strengths and positions	Aggregated vortex model	Pressure on the surface
Present Study[34, 35]	DA	EnKF	Incompressible 2D flow past a cylinder, a flat plate or an airfoil		100/200	Vorticity field and additional parameters	Unresolved DNS + Bias modeling	Velocity in the wake or pressure on the surface

1.3 Mathematical Background

By the very nature of an estimation problem, the state $x \in \mathbb{R}^n$ of any dynamical system[6] is only knowable up to a certain level of uncertainty. Mathematically, this behavior is represented by a random vector whose possible values correspond to individual realizations of the underlying random process. The probability that its value falls in any given subset $A \in \mathbb{R}^n$ of the state space is given by

$$\mathbb{P}(x \in A) = \int_A \rho(x)dx \,, \tag{1.1}$$

where the probability density function (PDF) $\rho : \mathbb{R}^n \mapsto \mathbb{R}^+$ satisfies

$$\int_{\mathbb{R}^n} \rho(x)dx = 1 \,. \tag{1.2}$$

For any function f of the state, we define the corresponding expected value by

$$\mathbb{E}[f(x)] = \int_{\mathbb{R}^n} f(x)\rho(x)dx \,. \tag{1.3}$$

When it is necessary to make explicit which PDF the expectation refers to, the notation $\mathbb{E}^\rho[\cdot]$ will be used.

An alternative way of describing the PDF of a random variable is through a sequence of central moments α_i given by

$$\alpha_i = \begin{cases} \bar{x} = \mathbb{E}[x] & \text{i=1} \\ \left. \frac{d^i M_x}{dt^i} \right|_{t=0} & \text{i>1} \end{cases} \,, \tag{1.4}$$

where $M_x(t) = \mathbb{E}[\exp(t^T(x - \bar{x}))]$[7] is the corresponding moment-generating function. In particular, note that the first-order moment $\alpha_1 = \bar{x}$ corresponds to the mean of the distribution, and the second-order central moment α_2 correspond to the auto-covariance matrix C^{xx}. According to the inverse theorem, if $M_x(t)$ is finite for all $t \in \{x \in \mathbb{R} \mid \|x\| < a\}$ for some $a > 0$, then $M_x(t)$ uniquely determines the PDF of x. Therefore, in order to track the time evolution of the state of a system, one needs to predict how the PDF changes over time, either directly or indirectly (by tracking its moments).

---

[4]IP stands for Inverse Problem, and DA stands for Data Assimilation. The latter differs from the former by assuming a temporal evolution of the variable being tracked.

[5]This quantity is defined differently for each problem.

[6]Despite the fact that the dynamics of fluid systems are inherently infinite-dimensional, we employ a discretization scheme to produce a finite-dimensional approximation suitable to be used in a computed simulation.

[7]Note that variable t here has the same dimension as x, so that $M_x(t)$ is a scalar function.

Besides the mean, another important feature of the PDF is the mode. The mode represents the most likely value of the corresponding random variable. Mathematically, it is given by

$$m_x = \arg\max_{x \in \mathbb{R}^n} \rho(x) \,. \tag{1.5}$$

Note that the mode need not to be unique and secondary modes corresponding to local maxima may be present.

1.3.1 Weighted Inner Product and Norm

For an ordered pair $u, v \in \mathbb{R}^n$, we denote the corresponding Euclidean inner product as

$$\langle u, v \rangle = u^T v \tag{1.6}$$

and its induced L2-norm as

$$\|u\|^2 = \langle u, u \rangle \geq 0 \,, \tag{1.7}$$

where the equality holds if, and only if, $u = 0$.

For any positive-definite symmetric matrix A, the A-induced weighted norm is given by

$$\|u\|_A = \|A^{-1/2}u\| \,. \tag{1.8}$$

1.3.2 Gaussian Random Variables

A Gaussian[8] random variable (GRV) on $\mathbb{R}^n$ is characterized by its mean $\bar{x} \in \mathbb{R}^n$ and a positive-semidefinite covariance matrix $C \in \mathbb{R}^{n \times n}$, and is often denoted as $x \sim N(\bar{x}, C)$. Its associated PDF is given by

$$\rho(x) = \frac{1}{(2\pi)^{n/2}(\det C)^{1/2}} \exp\left(-\frac{1}{2} \|x - \bar{x}\|_C^2\right) \,. \tag{1.9}$$

The matrix $C^{-1/2}$ is often called the precision matrix of the Gaussian random variable. Note also that for a GRV, the mean and mode of the distribution coincide. Gaussian-distributed random variables can be completely represented by their mean and variance, linear combination of GRVs are also Gaussian-distributed, and applying linear operators to such variables will always result in new Gaussian-distributed random variables.

---

[8]Sometimes referred to as a normal variable.

1.3.3 Conditional Distributions

If $(x, y) \in \mathbb{R}^{n \times p}$ is a jointly varying random variable, the conditional PDF $\rho(x|y)$ is defined as

$$\rho(x|y) = \frac{\rho(x, y)}{\rho(y)} , \tag{1.10}$$

where $\rho(y)$ corresponds to the marginal PDF of y given by

$$\rho(y) = \int_{\mathbb{R}^n} \rho(x, y) dx . \tag{1.11}$$

1.3.4 Bayes' Theorem

Bayes' theorem states that for a jointly varying random variable $(x, y) \in \mathbb{R}^{n \times p}$

$$\rho(x|y) = \frac{\rho(y|x)\rho(x)}{\rho(y)} . \tag{1.12}$$

This theorem is one of the cornerstones of the Bayesian inference, and expresses how prior beliefs $(\rho(x))$ should be updated to account for new evidence $(\rho(y))$ assuming a model that describes the likelihood of obtaining determined data given that the state $(\rho(y|x))$ is known. The resulting distribution $(\rho(x|y))$ is often referred to as the posterior.

1.3.5 Maximum a Posteriori vs Minimum Variance Estimates

An estimator $\tilde{x}(y)$ is a function of the observation $y \in \mathbb{R}^p$ that aims to provide the best estimate of x according to some criteria. In the Bayesian context, a first important notion of optimality is the minimization of the mean square error (MSE), given by the trace of the error covariance matrix

$$MSE(\tilde{x}) = tr \left\{ \mathbb{E}[(\tilde{x} - x)(\tilde{x} - x)^T] \right\} = \mathbb{E}[(\tilde{x} - x)^T(\tilde{x} - x)] . \tag{1.13}$$

The minimum-MSE estimator is then defined as the function

$$\tilde{x}_{MMSE}(y) = \arg \min_{\tilde{x}:\mathbb{R}^p \mapsto \mathbb{R}^n} MSE(\tilde{x}) . \tag{1.14}$$

It can be shown [36] that as long as the underlying PDF admits finite mean and covariance matrix, the minimum variance estimate is the conditional mean

$$\tilde{x}_{MMSE}(y) = \mathbb{E}[x|y] . \tag{1.15}$$

A second possible optimality criteria is carried out by the maximum a posteriori (MAP) estimator. It is defined as the mode of the posterior distribution obtained by applying Bayes' theorem.

$$\tilde{x}_{MAP}(y) = \arg\max_{\tilde{x}:\mathbb{R}^p \mapsto \mathbb{R}^n} \rho(\tilde{x}|y) = \arg\max_{\tilde{x}:\mathbb{R}^p \mapsto \mathbb{R}^n} \frac{\rho(y|\tilde{x})\rho(\tilde{x})}{\rho(y)}$$
$$= \arg\max_{\tilde{x}:\mathbb{R}^p \mapsto \mathbb{R}^n} \rho(y|\tilde{x})\rho(\tilde{x}) . \tag{1.16}$$

Note that when the posterior is Gaussian, the MMSE and the MAP estimator are equivalent since the mode and the mean of a Gaussian coincide.

A third criteria would be the maximum likelihood (ML) estimator. Here $\tilde{x}(y)$ is defined as the state that is more likely to produce the observed y, regardless of any a priori information on x. In other words,

$$\tilde{x}_{ML}(y) = \arg\max_{\tilde{x}:\mathbb{R}^p \mapsto \mathbb{R}^n} \rho(y|\tilde{x}) . \tag{1.17}$$

Note that the ML and MAP estimators coincide when the prior is a uniform distribution.

MMSE estimators are global in the sense that they summarize the information contained in the whole PDF when evaluating the conditional mean. On the other hand, MAP estimators only highlight local features of the posterior PDF, something that can be troublesome if the posterior is multi-modal, for instance.

1.3.6 Dynamical Systems

A general discrete-time controlled dynamical system can be characterized by a sequence of functions $f_k \in C(\mathbb{R}^n \times \mathbb{R}^c, \mathbb{R}^n)$, often referred to as the forecast model, that describe the time evolution of the state of the system subjected to some control input. These functions represent the available deterministic knowledge about the system at hand. When nondeterministic or unmodeled aspects of the underlying physical phenomenon are present, their effect can be taken into account stochastically. Consider the stochastic dynamical system

$$x_{k+1} = f_k(x_k, u_k) + \mu_k , \tag{1.18}$$

where f_k is a general function of the n-dimensional state x_k, u_k is a c-dimensional control input vector, and μ_k is an i.i.d. sequence with PDF ρ_s that accounts for state disturbances and process noises. The subscript k refers to quantities taken at time $t = t_k$. The trajectory of this system corresponds to a Markov chain for which

$$\rho(x_{k+1}|x_k) = \rho_s(x_{k+1} - f_k(x_k, u_k)) , \tag{1.19}$$

so that the probability that $x_{k+1} \in S \subset \mathbb{R}^n$ is given by

$$\mathbb{P}(x_{k+1} \in S|x_k) = \int_S \rho_s(x - f_k(x_k, u_k))dx \ . \tag{1.20}$$

We also assume that system is at least partially observable, and that there is a function $h \in C^1(\mathbb{R}^n, \mathbb{R}^p)$ that maps the state to any observable quantity. Since in any realistic measurement methodology involves uncertainties (noise), the typical observation model will take the form

$$y_k = h(x_k) + v_k \ . \tag{1.21}$$

where $h(x)$ is the observation function (y_k is a p-dimensional vector), and v_k is an i.i.d. sequence with PDF ρ_o that represents the sensor noise. Thus,

$$\rho(y_k|x_k) = \rho_o(y_k - h(x_k)) \ , \tag{1.22}$$

so that the probability that $y_k \in S \subset \mathbb{R}^p$ is given by

$$\mathbb{P}(y_k \in S|x_k) = \int_S \rho_o(y - h(x_k))dx \ . \tag{1.23}$$

1.4   Data Assimilation

Given a dynamical system, there are two ways of estimating its current state. A first one requires knowledge of the state at previous times, and a forecast model that describes the time evolution of the system. Combined, they can be used to estimate the trajectory of the system. The accuracy of theses predictions, however, relies not only on the preciseness of our estimate of the initial conditions but also on the reliability of the model itself. The fidelity of the chosen model is limited not only by our understanding of the dynamics of the system, but also by how fast are we expected to produce such estimate given computational resources.

A second approach requires access to measurement data from the actual system and an observation model, a function that approximately describes the mapping between the state of the system and any available observable quantity. The problem of finding the "best" estimate of the state(s) that conforms to the observed data according to the model at hand is an *inverse problem*. Again, the accuracy of this estimate relies upon the reliability of the assumed mapping between the state and measurements, and the signal-to-noise ratio of the available data.

Given these two estimates for the state of the system, one predicted from inaccurate previous estimates and another inferred from limited available data, the goal of the data assimilation is to determine the optimal strategy of managing the available resources and combining them in order to produce an estimate that meets the accuracy requirements.

DA methods can be classified as variational and sequential. The goal of variational methods is to optimize the estimated trajectory of the system over a given time interval while fulfilling model dynamics and conforming to the available measurements. The problem formulation involves the definition of a cost function (which penalizes the mismatch between predicted and measured data and enforces the system dynamics) in terms of a control variable (e.g., initial state). The evaluation of the optimal control variable typically involves iteration and derivatives in the form of a sensitivity, or adjoint, model. On the other hand, sequential methods get their name from the fact that they are usually formulated as a sequential repetition of two basic steps: when new measurements are available, corrections are applied to the state estimate (analysis step); then, the state statistics are propagated forward in time using the available dynamic model (forecast step) until new measurements become available again. Figure 1.2 shows a graphical representation of a sequential estimation process. The points and oval regions represent, respectively, the mean and uncertainty (which can be seen as a contour level of the underlaying PDF) of each estimate. Estimation methodology combines model prediction (blue figure) to the state inferred from the measurements (red figure) to yield an improved state estimate (green figure).

Algorithmically speaking, sequential methods are more flexible with respect to the choice of forecast models. Because they do not require the corresponding linearized forward and adjoint operators, models can be treated as black boxes. In terms of performance, sequential methods tend to favor the use of more complex models in real-time applications, since a single forward integration is needed.

In the following sections, we present a summary of the existing state of knowledge for sequential methods as it is presented in the literature[19, 36, 38–40].
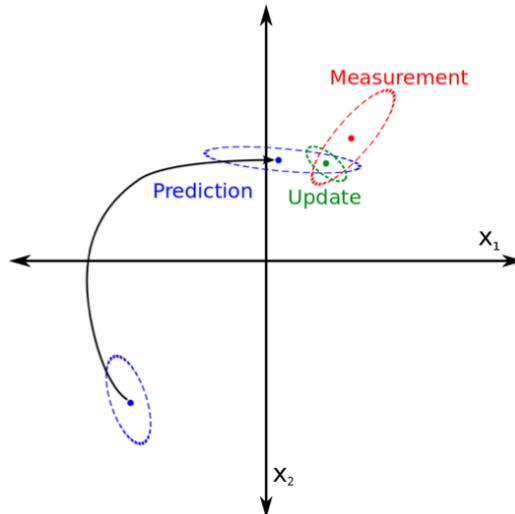
Figure 1.2: Graphical representation of the estimation process. Reproduced from [37] with permission.

1.4.1 Kalman Filtering

The fundamentals of optimal filtering[9] were laid down by Kalman [4]. The classical Kalman filter provides a rigorous solution for the state estimation of linear systems under Gaussian-distributed noise. The goal of Kalman filtering is to use measurement data to construct an estimate of the state x_k which is optimal in the sense that it minimizes the estimation variance (or, equivalently, maximizes the likelihood) [41]. The estimate is regarded as a Gaussian-distributed random variable which is characterized by its mean x_k and covariance C_k. Assuming linearity, $f_k(x, u) = F_k x + B_k u$ and $h(x) = Hx$ Eq. 1.18 and 1.21 can be rewritten as

$$x_{k+1} = F_k x_k + B_k u_k + \mu_k \tag{1.24a}$$

$$y_k = H_k x_k + v_k . \tag{1.24b}$$

We also assume both μ_k and v_k are zero-mean, Gaussian, and white random processes ($\mu_k \sim N(0, Q_k)$ and $v_k \sim N(0, R_k)$) that are uncorrelated in time ($E[\mu_k \mu_l^T] = Q_k \delta_{kl}$ and $E[v_k v_l^T] = R_k \delta_{kl}$, where δ_{kl} is the Kronecker delta)[10].

Defining $Y_k = \{y_1 \; y_2 \; \cdots \; y_k\}$ as the set that collects all the measurements taken from the system up to time $t = t_k$, it can be shown that the optimum filtering process for the unforced system ($u = 0$) can be synthesized in two steps [42]:

---

[9]In this context, filtering is used to refer to the problem of determining the state of a system from noisy measurements.

[10]The matrices $R_k \in \mathbb{R}^{p \times p}$ and $Q_k \in \mathbb{R}^{n \times n}$ are called the covariance matrices for the measurement and process noises, respectively.

- **Dynamic update (or Forecast Step):** The mean and covariance of the state at the next assimilation step is estimated using Eq. 1.24a.

  If the initial state is represented by a Gaussian random vector with mean $\bar{x}_0$ and C_0, the linear dynamics will preserve Gaussianity and the forecast is completely described by its mean and covariance. Since the noise is independent of the state

$$
\begin{aligned}
\bar{\hat{x}}_{k+1} &= \mathbb{E}[x_{k+1}|Y_k] = \mathbb{E}[F_k x_k|Y_k] + \mathbb{E}[\mu_k|Y_k] \\
&= F_k \mathbb{E}[x_k|Y_k] = F_k \bar{x}_k \tag{1.25a} \\
\hat{C}_{k+1} &= \mathbb{E}[(x_{k+1} - \bar{\hat{x}}_{k+1})(x_{k+1} - \bar{\hat{x}}_{k+1})^T|Y_k] \\
&= \mathbb{E}[F_k(x_k - \bar{\hat{x}}_k)(x_k - \bar{\hat{x}}_k)^T F_k^T|Y_k] + \mathbb{E}[\mu_k \mu_k^T|Y_k] \\
&\quad + \mathbb{E}[\mu_k(x_k - \bar{\hat{x}}_k)^T F_k^T|Y_k] + \mathbb{E}[F_k(x_k - \bar{\hat{x}}_k)\mu_k^T|Y_k] \\
&= F_k C_k F_k^T + Q_k \ , \tag{1.25b}
\end{aligned}
$$

  where the hat is used to represent forecast variables.

- **Measurement update (or Analysis Step):** A new set of measurement data is incorporated into the estimate following Bayes' rule. If the prior distribution corresponds to the forecast estimate, the posterior distribution is given by

$$
\rho(x_{k+1}|Y_{k+1}) = \frac{\rho(y_{k+1}|x_{k+1}, Y_k)\rho(x_{k+1}|Y_k)}{\rho(y_{k+1}|Y_k)} \ . \tag{1.26}
$$

  Since Gaussian distributions are self-conjugate with respect to Gaussian likelihoods, the posterior distribution is also Gaussian:

$$
\exp\left(-\frac{1}{2}\|x - \bar{x}_{k+1}\|^2_{C_{k+1}}\right) = \alpha \exp\left(-\frac{1}{2}\|y_{k+1} - H_{k+1}x\|^2_R - \frac{1}{2}\|x - \bar{\hat{x}}_{k+1}\|^2_{\hat{C}_{k+1}}\right) , \tag{1.27}
$$

  where α is a normalizing constant. Equating quadratic and linear terms in x in the exponents gives, respectively

$$
C_{k+1}^{-1} = \hat{C}_{k+1}^{-1} + H_{k+1}^T R^{-1} H_{k+1} \tag{1.28a}
$$

$$
C_{k+1}^{-1}\bar{x}_{k+1} = \hat{C}_{k+1}^{-1}\bar{\hat{x}}_{k+1} + H_{k+1}^T R^{-1} y_{k+1} \ . \tag{1.28b}
$$

  Using the Woodbury matrix identity, Eq. 1.28a becomes

$$
\begin{aligned}
C_{k+1} &= \left(\hat{C}_{k+1}^{-1} + H_{k+1}^T R^{-1} H_{k+1}\right)^{-1} \\
&= \hat{C}_{k+1} - \hat{C}_{k+1} H_{k+1}^T \left(R + H_{k+1}\hat{C}_{k+1} H_{k+1}^T\right)^{-1} H_{k+1}\hat{C}_{k+1} \\
&= (I - K_{k+1} H_{k+1})\hat{C}_{k+1} \ , \tag{1.29}
\end{aligned}
$$

where $K_{k+1} = \hat{C}_{k+1}H_{k+1}\left(R + H_{k+1}\hat{C}_{k+1}H_{k+1}^T\right)^{-1}$ is the so called Kalman gain. Note that here the inversion is performed in the measurement space (p-by-p matrix inversion). The Woodbury matrix identity can be used again to yield

$$
\begin{aligned}
K_{k+1} &= \hat{C}_{k+1}H_{k+1}\left(R + H_{k+1}\hat{C}_{k+1}H_{k+1}^T\right)^{-1} \\
&= \hat{C}_{k+1}H_{k+1}\left(R + H_{k+1}\hat{C}_{k+1}H_{k+1}^T\right)^{-1} \times \\
&\qquad\qquad \left(\left(R + H_{k+1}\hat{C}_{k+1}H_{k+1}^T\right) - H_{k+1}\hat{C}_{k+1}H_{k+1}^T\right) R^{-1} \\
&= \hat{C}_{k+1}H_{k+1}\left(I - \left(R + H_{k+1}\hat{C}_{k+1}H_{k+1}^T\right)^{-1} H_{k+1}\hat{C}_{k+1}H_{k+1}^T\right) R^{-1} \\
&= \left(\hat{C}_{k+1} - \hat{C}_{k+1}H_{k+1}^T\left(R + H_{k+1}\hat{C}_{k+1}H_{k+1}^T\right)^{-1} H_{k+1}\hat{C}_{k+1}\right) H_{k+1}^T R^{-1} \\
&= \left(\hat{C}_{k+1}^{-1} + H_{k+1}^T R^{-1}H_{k+1}\right)^{-1} H_{k+1}^T R^{-1} ,
\end{aligned}
\tag{1.30}
$$

which corresponds to a inversion in the state space (n-by-n matrix inversion).

Substituting Eq. 1.29 in Eq. 1.28b,

$$
\begin{aligned}
\bar{x}_{k+1} &= C_{k+1}\left(\hat{C}_{k+1}^{-1}\bar{\hat{x}}_{k+1} + H_{k+1}^T R^{-1}y_{k+1}\right) \\
&= (I - K_{k+1}H_{k+1})\bar{\hat{x}}_{k+1} + \left(\hat{C}_{k+1}^{-1} + H_{k+1}^T R^{-1}H_{k+1}\right)^{-1} H^T R^{-1}y_{k+1} \\
&= (I - K_{k+1}H_{k+1})\bar{\hat{x}}_{k+1} + K_{k+1}y_{k+1} = \bar{\hat{x}}_{k+1} + K_{k+1}\left(y_{k+1} - H_{k+1}\bar{\hat{x}}_{k+1}\right) ,
\end{aligned}
\tag{1.31}
$$

where the difference $\left(y_{k+1} - H_{k+1}\bar{\hat{x}}_{k+1}\right)$ is often referred to as the innovation vector. If the Kalman filter works optimally, the innovation sequence is expected to be white.

Another important consequence of the fact that the posterior is a Gaussian is that its mode can be used as a proxy for its mean. Therefore,

$$
\begin{aligned}
\bar{x}_{k+1} &= \arg\max_{x\in\mathbb{R}^n} \exp\left(-\frac{1}{2}\|y_{k+1} - H_{k+1}x\|_R^2 - \frac{1}{2}\|x - \bar{\hat{x}}_{k+1}\|_{\hat{C}_{k+1}}^2\right) \\
&= \arg\min_{x\in\mathbb{R}^n} \left(\frac{1}{2}\|y_{k+1} - H_{k+1}x\|_R^2 + \frac{1}{2}\|x - \bar{\hat{x}}_{k+1}\|_{\hat{C}_{k+1}}^2\right) .
\end{aligned}
\tag{1.32}
$$

This last equation highlights the optimization perspective of the analysis step of the Kalman Filter[43], and will play a fundamental role later on in this work.

Taken together, these two steps resemble a Luenberger observer with an adaptive observer gain. It requires the storage and propagation of the covariance matrix, an operation that has a nominal complexity of $O(n^3)$. Therefore, the computational cost of the filter rapidly increases with the number of the degrees of freedom of the system and soon becomes intractable for real-time applications.

The Extended Kalman Filter (EKF)

Devising an optimum state estimator for systems modeled by nonlinear dynamics with measurement data that is a nonlinear function of the state is more challenging. Prospects for rigorously addressing the problem typically end up being too narrow in applicability or too computationally expensive [44].

For weakly nonlinear problems, the so-called Extended Kalman Filter (EKF) [45] is considered the standard technique. Assuming the dynamics are weakly nonlinear, the EFK linearizes the dynamics about the estimate mean and uses the resulting Jacobian matrices (Eq. 1.33) to evaluate the Kalman gain and update the surrogate covariance matrix. In most cases, the nonlinear dynamics is still used to update the estimate mean.

$$F_k = \left.\frac{\partial f}{\partial x}\right|_{x=\hat{x}} \qquad H_k = \left.\frac{\partial h}{\partial x}\right|_{x=\hat{x}} \tag{1.33}$$

Note that the EKF still requires an explicit evaluation of the covariance matrix and tracks its evolution as the simulation progresses. In comparison to the standard Kalman Filter, the EKF incurs the extra cost of computing the appropriate linearization at each assimilation step.

Higher-order Kalman Filters

According to nonlinear filter theory [36], in general, the evolution of the conditional mean and covariance matrix depends on all the moments (an infinite number of them) of the conditional density function. In fact, the time-evolution of the state PDF is governed by the Fokker-Planck equation (also known as the Kolmogorov Forward equation), a n-dimensional PDE. The only exception is, as shown by Jazwinski [36], the classical KF (Gaussian prior and likelihood, and linear forecast and observation models), for which the first two moments fully describe the filter.

Therefore, numerically, we must consider finite-size approximations of the posterior PDF by means of a parametrization. If all PDFs can be considered nearly-Gaussian,

all odd central moments can be neglected, and higher-order even moments can be written in terms of the variance. The extended KF corresponds to the case where all moments higher than second are neglected. A corresponding second-order filter can be obtained by retaining the 4th-order central moments. In that case, the resulting second-order filter is given by[36]

- **Forecast**:

$$\bar{\hat{x}}_{k+1} = \mathbb{E}[x_{k+1}|Y_k] = f_k(\bar{x}_k) + \frac{1}{2}\left[\partial^2 f \hat{C}_k\right] \tag{1.34a}$$

$$\hat{C}_{k+1} = \left[\frac{\partial h}{\partial x}(\hat{x}_k)\right] C_k \left[\frac{\partial h}{\partial x}(\hat{x}_k)\right]^T + Q_k . \tag{1.34b}$$

- **Analysis**:

$$\left(R + \left[\frac{\partial h}{\partial x}(\hat{x}_k)\right] \hat{C}_k \left[\frac{\partial h}{\partial x}(\hat{x}_k)\right]^T + \frac{1}{2}\left[\partial^2 h \hat{C}_k^2 \partial^2 h\right]\right) b_k^*$$

$$= y_k - \mathbb{E}[h(x_k)|Y_{k-1}] - \frac{1}{2}\left[\partial^2 h \hat{C}_k\right] \tag{1.35a}$$

$$z_k = \hat{z}_k + \hat{C}_k \left[\frac{\partial h}{\partial x}(\hat{x}_k)\right]^T b_k^* , \tag{1.35b}$$

where

$$\left\{\partial^2 h \hat{C}_k^2 \partial^2 h\right\}_{ij} = \sum_{k,l,p,q=1}^{n} \frac{\partial^2 h_i}{\partial x_k \partial x_l}(\hat{x}_k) \frac{\partial^2 h_j}{\partial x_p \partial x_q}(\hat{x}_k) \left\{\hat{C}_k\right\}_{lp} \left\{\hat{C}_k\right\}_{kq} \tag{1.36a}$$

$$\left\{\partial^2 h \hat{C}_k\right\}_i = \sum_{j,k=1}^{n} \left\{\hat{C}_k\right\}_{jk} \frac{\partial^2 h_i}{\partial x_j \partial x_k}(\hat{x}_k) . \tag{1.36b}$$

Note that the second-order filter equation differs from the EKF equations by the presence of the terms $\left[\partial^2 f \hat{C}_k\right]$, $\left[\partial^2 h \hat{C}_k\right]$ and $\left[\partial^2 h \hat{C}_k^2 \partial^2 h\right]$. This implies that the effect of nonlinearities in the dynamics of the conditional mean depends on the magnitude of the product between the second partial derivative of the forecast and observation models, and the estimation error variance. Therefore, even for systems in which the second partial derivatives of the model are small in comparison to their respective first derivatives, the nonlinearities can play a important role at early times of the estimation history, when the variance of the estimate is typically large. Moreover, analyzing Eq. 1.35a (see appendix D), we can note that the addition of the extra term to the RHS plays the same role of a bias correction term and ensures the innovation vector has zero expectation (an optimality feature of the Kalman

filter). Note that its magnitude scales with the estimation variance. Furthermore, the importance of the extra term added to the LHS can be accessed by comparing it to the magnitude of R. Because it has a damping effect on the corrections, ignoring it leads to overcorrections.

The Unscented Kalman Filter (UKF)

The robustness and reliability of the EKF is impaired by the linearization process. For example, Julier and Uhlmann [46] showed that even a trivial nonlinear transformation from polar to Cartesian coordinates is enough to yield significant deviations in tracking the correct state. For cases where there are strong nonlinearities, the Unscented Kalman Filter tends to deliver better results [47]. This scheme provides an alternative to the explicit evaluation of the second derivatives of the forecast and observation model, by employing a deterministic sampling scheme (unscented transformation) to generate a set of points around the prior mean (sigma points) which are propagated by the nonlinear functions and then used to reconstruct the posterior mean and variance with second-order accuracy. Although it has been demonstrated that it delivers excellent results, it requires $4n + 1$ sigma points (or $2n + 1$, if there is no process noise) for the forecast step and $4p + 1$ sigma points for the analysis step, and the corresponding computational cost[11] becomes prohibitive for large systems.

1.4.2 3D-Var

A well-known alternative for sequential data assimilation of high-dimensional systems is the 3D-Var[48]. Just like the Kalman filter, 3D-Var can be formulated as an observer in which the gain is calculated to minimize a cost function, with the general format

$$J(x) = [y_0 - h(x)]^T R^{-1} [y_0 - h(x)] + \left[x - x_f\right] \Sigma^{-1} \left[x - x_f\right] \, , \qquad (1.37)$$

where y_0 is the measurement taken from the tracked system, x_f is the prior estimate for the state, $h(x)$ is the observation function. R, as a measure of the reliability of the measurements, is a constant matrix that weights the measurement mismatch. Differently from the KF methodology that regards Σ as time-varying estimate of the

---

[11]Although more expensive than a KF for a given number of degree of freedom, the cost scaling is the same ($O(n^3)$) if we assume the forecast model to have cost $O(n^2)$.

state covariance matrix, 3D-Var regards it as a predefined constant weight matrix. Since there is no explicit tracking of the covariance matrix, 3D-Var is far less computationally demanding than KF, but estimator performance depends heavily on the *a priori* choice of Σ. For a general function $h(x)$, the minimizer of Eq. 1.37 is usually obtained using an appropriate iterative method (e.g., quasi-Newton method). If $h(x) = Hx$ is a linear function, the minimizer of the aforementioned cost function is given by

$$x_k = \bar{x}_k + K(y_k - H\bar{x}_k),$$ (1.38)

where $K = \Sigma H^T \left(R + H\Sigma H^T \right)^{-1}$ is a constant matrix.

1.4.3 Ensemble Methods

The most complete description of the state of the system under scrutiny is given by its probability density function $\rho(x)$. For applications in fluid mechanics, the dimension of the domain of integration can easily be of order $O(10^6)$ or more, and a direct numerical evaluation of this integral (using an appropriate quadrature rule) is prohibitive. This difficulty can be interpreted as one of the manifestations of the *curse of dimensionality*, first introduced by Bellman [49]. It indicates that the number of samples needed to estimate an arbitrary function with a given level of accuracy grows exponentially with respect to the number of degrees of freedom (i.e., dimensionality) of the function.

Instead, one can adopt a Monte Carlo approach and represent the PDF as a combination of Dirac delta functions corresponding to an ensemble of q independently drawn random samples from $\rho(x)$, denoted as x_i,

$$\rho_{MC}(x) = \frac{1}{q} \sum_{i=1}^{q} \delta(x - x_i),$$ (1.39)

for which the expected value of $f(x)$ can be evaluated as

$$\mathbb{E}_{MC}[f(x)] = \int f(x)\rho_{MC}(x)dx = \frac{1}{q} \sum_{i=1}^{q} f(x_i).$$ (1.40)

It can be shown [39] that $\mathbb{E}_{MC}[f(x)]$ is an unbiased estimate of $\mathbb{E}[f(x)]$, whose variance converges to zero according to

$$Var[\mathbb{E}_{MC}[f(x)]] = \frac{1}{q}Var[f(x)],$$ (1.41)

where

$$Var[f(x)] = \frac{1}{q-1} \sum_{i=1}^{q} (f(x_i) - \mathbb{E}[f(x)])^2 \,. \tag{1.42}$$

Note that rate of convergence $1/\sqrt{q}$ is independent of the dimensionality of the state. That notwithstanding, the number of particles required to achieve a given accuracy threshold surely depends on the dimensionality of the state. Usually, one can expect at least some number in the same order of magnitude of the dimensionality of the state.

For all the KF variants that were discussed so far, the computational cost of the propagation of the covariance matrix C_k has the same order of magnitude as n evaluations of the forecast operator and becomes quickly prohibitive for large systems.

Particle Filter (PF)

Under nonlinear dynamics, the Probability Density Function (PDF) of the estimate need not remain Gaussian, and the first two moments cease to fully represent the state. Instead, the Fokker-Plank PDE, that describes the evolution of the full PDF in time, can be discretized using a Lagrangian method (something that can be interpreted as a Monte Carlo sampling) to yield what is commonly referred to as Particle filters[23, 50]. Here no assumption is made on the shape of the state PDF. The dynamical model is responsible for forecasting the trajectory of these particles through time. Measurement data is incorporated into the description of the PDF by assigning weights to each of the particles which are updated whenever new measurements are available. Thus,

$$\rho_{MC}(x|y) = \frac{1}{q} \sum_{i=1}^{q} w_i \delta(x - x_i) \,, \tag{1.43}$$

where the weights are computed as normalized likelihoods:

$$w_i = \frac{\rho(y|x_i)}{\sum_{i=1}^{q} \rho(y|x_i)} \,. \tag{1.44}$$

The PDF of the observations conditioned to the state is described by measurements statistics, and is usually taken to be a Gaussian:

$$\rho(y|x_i) \sim \exp\left(-\frac{1}{2} [y - h(x_i)]^T R^{-1} [y - h(x_i)]\right) \,. \tag{1.45}$$

Note that the measurements don't influence the trajectory of the particle. Therefore, there is no guarantee that the particles will remain in the region of the state space that is relevant to the measurement data obtained. Consequently, a considerable fraction of the ensemble may end up with negligible weights, hindering the accuracy of the scheme. This feature of the method is the most common cause of filter divergence. In order to avoid this issue, the ensemble must be constantly resampled to ensure the particle to remain relevant. Leeuwen [50] describe several methodologies to accomplish this task. That notwithstanding, because PF schemes rely on the direct sampling of a n-dimensional state space, the curse of dimensionality [49] is severe for these techniques, and they are only computationally tractable for systems of reduced dimension.

If the state and the likelihood can be described as approximately Gaussians, the need for resampling can be eliminated by employing an ensemble variant of the Kalman filter, namely the Ensemble Kalman Filter (EnKF).

1.5   The Ensemble Kalman Filter (EnKF)

Aiming at overcoming the computational cost limitation, Evensen [51] proposed a Monte Carlo approximation to the KF in which the internal state of the estimator is represented by an ensemble of particles so that the corresponding ensemble mean and covariance matrix are used to approximate $\bar{x}$ and C. This method was named Ensemble Kalman Filter (EnKF), and since then has been extensively used for high-dimensional systems (thousands of degrees of freedom or more) associated with a computationally onerous forecast (as in meteorology, oceanography and geophysical flows Bengtsson, Snyder, and Nychka [52], Evensen [53], and Anderson and Anderson [54]). In such context, this technique has shown to provide a correct estimate of the first two moments of state of the system even for a small ensemble size (provided that the Gaussian assumption appears to hold) [55].

The main advantages of the EnKF in comparison to the variational methods or standard KF formulations are:

- It does not require the adjoint of the dynamical model.

- It has low storage requirement (comparing to the storage needed to store the full state statistics).

- It has a natural probabilistic interpretation under a Bayesian perspective.
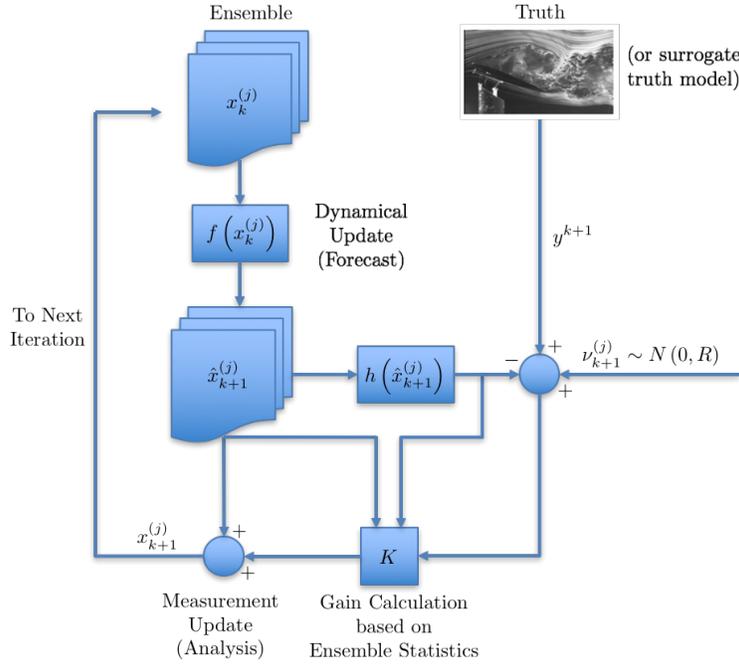
Figure 1.3: Schematic representation of the EnKF algorithm.

- Due to its formulation in terms of independent particles, it is embarassingly parallel.

1.5.1 Algorithm

Figure 1.3 shows a schematic diagram of the EnKF algorithm. Having an ensemble-based representation of the state in mind, consider an ensemble of q initially independent states sampled from a normal distribution with predefined mean and covariance matrix. The expected value of the system state corresponds to the ensemble average of these states

$$\bar{x}_k = \frac{1}{q} \sum_{j=1}^{q} x_k^{(j)} \,. \tag{1.46}$$

Defining the scaled state perturbation matrix A_k as

$$A_k = \frac{1}{\sqrt{q-1}} [x_k^{(1)} - \bar{x}_k \cdots x_k^{(q)} - \bar{x}_k] \,, \tag{1.47}$$

and the scaled output perturbation matrix HA_k (assuming the linearity of the observation function, i.e., $h(x) = Hx$) as

$$HA_k = \frac{1}{\sqrt{q-1}} [y_k^{(1)} - \bar{y}_k \cdots y_k^{(q)} - \bar{y}_k] \,, \tag{1.48}$$

where $y_k^{(j)} = h(x_k^{(j)})$ and $\bar{y}_k$ is the ensemble average of the outputs, one can finally compute the ensemble covariance matrix

$$C_k = A_k(A_k)^T , \tag{1.49}$$

which is an estimate of the state covariance matrix.

The filtering algorithm can be summarized as a succession of two steps: a forecast step (or dynamic update) and a analysis step (or measurement update).

Forecast Step

The state of each ensemble member at the next time step is estimated using the (possibly nonlinear) dynamic model (Eq. 1.18):

$$\hat{x}_{k+1}^{(j)} = f(x_k^{(j)}, u_k) + g(x_k^{(j)})\mu_k^{(j)} , \tag{1.50}$$

where the hat is used to represent forecast variables. If applied to a linear system, this ensemble approach reduces the cost associated with the time-propagation of the state statistics from $O(n^3)$ (classical KF) to $O(n^2 q)$ (EnKF). Since typical ensemble sizes are no larger than $O(10^2)$, the overall cost is usually reduced by several orders of magnitude.

Analysis Step

The ensemble members are corrected in order to minimize the error with respect to the measurements in the presence of noise and model uncertainties. There are several paths that lead to the EnKF analysis formula. We here adopt the optimization approach by looking for the minimizer of the cost function

$$
\begin{aligned}
J(x) &= \frac{1}{2} \|y_k - Hx\|_R^2 + \frac{1}{2} \|x - \hat{x}_k\|_{\hat{C}_k}^2 \\
&= \frac{1}{2}[y_k - Hx]^T R^{-1}[y_k - Hx] + \frac{1}{2}[x - \hat{x}_k]^T \hat{C}_k^{-1}[x - \hat{x}_k] .
\end{aligned}
\tag{1.51}
$$

This optimization problem is then restricted to the affine space generated by the prior estimate of each of the ensemble members and the subspace spanned by the scaled perturbation matrix $\hat{A}_k$. In other words, we look for a solution in the form

$$x = \hat{x}_k + \hat{A}_k v , \tag{1.52}$$

where $v \in \mathbb{R}^q$ is the correction coefficient vector.

After performing the proposed change of variables, we can restate the objective of the analysis step as finding

$$v = \arg\min_{v \in \mathbb{R}^q} J(v) \tag{1.53}$$

for each of the ensemble members, where

$$J(v) = \frac{1}{2}\|v\|^2 + \frac{1}{2}\|y_k - H\hat{x}_k - H\hat{A}_k v\|_R^2 . \tag{1.54}$$

Since $J(v)$ is quadratic in v, the solution is unique and corresponds to the root of

$$DJ(v) = v - (H\hat{A}_k)^T R^{-1}(y_k - H\hat{x}_k - H\hat{A}_k v) = 0 , \tag{1.55}$$

which is given by

$$v_k = \left[I + (H\hat{A}_k)^T R^{-1}(H\hat{A}_k)\right]^{-1} (H\hat{A}_k)^T R^{-1}(y_k - H\hat{x}_k) \tag{1.56a}$$

$$= (H\hat{A}_k)^T \left[R + (H\hat{A}_k)(H\hat{A}_k)^T\right]^{-1} (y_k - H\hat{x}_k) , \tag{1.56b}$$

where we have used the Woodbury matrix identity to obtain the alternative solution.

Notice that here we have the possibility to choose between performing the analysis in the ensemble space (q-by-q matrix inversion - Eq. 1.57a), or in the measurement space (p-by-p matrix inversion - Eq. 1.57b), depending on which one is more advantageous[56]. In either case, $q \ll n$ or $p \ll n$ such that an enormous reduction in computational expense is achieved compared to the KF/EKF.

The final solution is then obtained by projecting these coefficients back to the state space:

$$x_k = \hat{x}_k + \hat{A}_k \left[I + (H\hat{A}_k)^T R^{-1}(H\hat{A}_k)\right]^{-1} (H\hat{A}_k)^T R^{-1}(y_k - H\hat{x}_k) \tag{1.57a}$$

$$= \hat{x}_k + \hat{A}_k(H\hat{A}_k)^T \left[R + (H\hat{A}_k)(H\hat{A}_k)^T\right]^{-1} (y_k - H\hat{x}_k) . \tag{1.57b}$$

Algorithmically, when the inversion is done in the measurement space, instead of solving for v_k, the representers' formulation proposed by Evensen and Leeuwen [57] is used:

$$\left[R + (H\hat{A}_k)(H\hat{A}_k)^T\right] b_k = (y_k - H\hat{x}_k) \tag{1.58a}$$

$$x_k = \hat{x}_k + \hat{A}_k(H\hat{A}_k)^T b_k , \tag{1.58b}$$

where the columns of $\hat{A}_k(H\hat{A}_k)^T$ are called the *representers* and represent the influence vectors for each measurement. The vector b_k then represents the magnitude by which each of the representers actuates in $\hat{x}$.

For each ensemble member, y_k must be independently sampled from a normal distribution whose mean is the measurement vector obtained from the estimated system, and whose variance is R_k. Due to this sampling step, this algorithm is often referred to as perturbed observations (or stochastic) EnKF. Although this procedure introduces an additional sampling error, previous work by Lawson and Hansen [58] suggested it performs better in the presence of nonlinearities than deterministic alternatives.

It is worthy to note that one never needs to explicitly compute the covariance $\hat{C}_k$ since it suffices to evaluate $\hat{A}_k(H\hat{A}_k)^T$ and $H\hat{A}_k(H\hat{A}_k)^T$. Both the Particle Filter (PF) and the EnKF algorithms share the same forecast step, but their analysis steps are distinct. While in the PF the posterior PDF corresponds to a linear combination of the prior ensemble whose weights are calculated using the Bayes' rule, the EnKF assign equal weights to all particles and correct the ensemble members themselves according to Kalman's update rule[23]. Because the particles themselves are driven towards the measurements, the need for resampling is eliminated.

1.5.2 Initialization Scheme

In order to keep the cost at tractable levels, it is desirable to use ensemble sizes that are much smaller than the dimension of the state. Thus, being able to efficiently sample the initial ensemble plays a fundamental role in the filter performance. Therefore, following Evensen's scheme[19]:

1. Using a long series of snapshots obtained from a long simulation of the phenomenon we are interested in, we build the data matrix $\hat{X} = [x^{(1)} \, x^{(2)} \, \cdots \, x^{(N)}]$ and obtain the corresponding POD modes by computing the singular value decomposition

$$\bar{x} = \frac{1}{N}\hat{X}\mathbb{1}_{N\times 1} \tag{1.59}$$

$$\frac{1}{\sqrt{N-1}}(\hat{X} - \bar{x}\mathbb{1}_{1\times N}) = U\Sigma V^T , \tag{1.60}$$

where $\bar{x}$ is the mean flow.

2. In order to generate q independent initial ensemble members, we restrict it to the subspace spanned by the first q POD modes:

$$A_0 = qr(randn(q)) \tag{1.61}$$

$$X_0 = \sqrt{q-1}\tilde{U}_q\tilde{\Sigma}_q A_0 \,, \tag{1.62}$$

where $\tilde{U}_q$ corresponds to the first q columns of U, and $\tilde{\Sigma}_q$ is the upper-left $q \times q$ submatrix of Σ. Here $randn(q)$ represents a $q \times q$ matrix whose entries where independently sampled from a zero-mean Gaussian distribution with unitary variance, and $qr(\cdot)$ correspond to an implementation of the QR decomposition.

A similar approach can be used to generate the ensemble of noise vectors needed to force the dynamics:

$$A_k = qr(randn(q)) \tag{1.63}$$

$$M_k = \alpha\sqrt{q-1}\tilde{U}_q\tilde{\Sigma}_q A_k \,, \tag{1.64}$$

where α is a parameter that controls the noise magnitude. In this case, the corresponding error covariance matrix is given by $Q = \alpha^2\tilde{U}_q\tilde{\Sigma}_q^2\tilde{U}_q^T$.

1.5.3 Ensemble Size

Were the KF hypotheses to hold (linearity of forecast and observation models and Gaussianity of all variables and noise processes), optimality can only be expected as $q \to \infty$. Under nonlinear dynamics, this sub-optimum filter can only be expected to provide estimates for the first two moments of a possible general state PDF. In fact, Le Gland, Monbet, and Tran [59] demonstrated that in general the EnKF exhibits a $q^{-1/2}$ rate of convergence to the estimate asymptotic PDF. However, this limiting distribution may differ from the optimum filtering distribution, which may be multi-modal or exhibit other higher-order features in the presence on nonlinearities.

In a practical setting, the ensemble size q required to guarantee accuracy will depend on the effective state space dimension (the dimensionality of the manifold in which the most relevant dynamics is confined) and on our ability to recognize and sample this subspace[55]. Nevertheless, typical ensemble sizes do not exceed a few hundreds in view of the available computational power, and for such small ensemble sizes the EnKF exhibits some interesting characteristics.

Since the correction added to each ensemble member during the analysis step is a combination of the forecast ensemble states, the analysis step operates only in the subspace spanned by the ensemble at hand [60]. Therefore, the choice of initial ensemble can limit, at least initially, the effectiveness of the correction applied by the estimator to the ensemble members. The impact of this choice on the long-term behavior of the filter is highly dependent on the dynamical system itself. A very strong attractor, for instance, can render the choice of initialization scheme secondary.

1.5.4 Covariance Inflation

Many sources of error in the EnKF are extrinsic, i.e., external, to the filtering scheme being used. These include the loss of Gaussianity due to the presence of nonlinearities and by inaccurate models. Dealing with these errors will be a major theme of this thesis, but they all have in common the fact that they are related to the choice of forecast and observation model and are not directly related to the EnKF algorithm.

On the other hand, some error sources are intrinsic to the EnKF methodology and may be present even when perfect linear models are employed and all Gaussianity requirements are met. The main source of intrinsic error in the ensemble Kalman filter is the sampling error [61]. As pointed out in the previous section, even though the ensemble statistics are expected to slowly converge to their true value, the improved performance claim laid out by ensemble methods is built upon the hypothesis that just a reduced number of ensemble members ($q < 100$) may suffice to represent the first and second-order statistics of high-dimensional systems ($n \gg$).

The undesirable effects off under-sampling are twofold: generation of spurious cross-correlations, and underestimation of error covariances. When Houtekamer and Mitchell [62] first reported these issues, they used the analogy of *inbreeding* to describe the problem: the gain used to update the ensemble is computed using estimated for the statistics evaluated from the same ensemble. Later, Leeuwen [63] noted that a finite ensemble size leads to a consistent underestimation of the error variance. To use a CFD analogy, the EnKF behaves as if the error term in the modified equation (where the sampling takes the place of the discretization) always acts by damping the variance of the system.

In the absence of process noise and for a fixed measurement noise level, as the estimated covariance decreases, the weight given to the measurement data in the

analysis step decreases and eventually becomes negligible. This phenomenon is known as *covariance collapse*, in reference to the fact that the ensemble members collapse onto a single trajectory. The effect of the sampling errors will accelerate this collapse, and this leaves the filter vulnerable to ignoring new information contained in the measurements.

Aiming to mitigate the effects of the under-sampling errors, Bocquet [61, 64, 65] proposed a variant of the EnKF called Finite-size Ensemble Kalman Filter (EnKF-N) that assumes a different prior in the analysis step. Since the true values of the ensemble mean and covariance matrix are unknown, Bocquet argues that the prior should be represented by a multivariate Student's t-distribution with $q - 1$ degrees of freedom. The associated cost function

$$J(v) = \frac{1}{2} \left\| y - h(\hat{x}_k + \hat{A}_k v) \right\|_R^2 + \frac{q}{2} \ln \left(1 + \frac{1}{q} + \|v\|^2 \right) \tag{1.65}$$

no longer has an explicit analytical formula for its minimizer (even when $h(x)$ is linear), and the solution needs to be found iteratively. A nice feature of this formulation is that the Gaussian prior is recovered as the ensemble size tends to infinity. Even though this scheme obtained good results when compared to other methodologies to mitigate under-sampling errors, the cost involved in minimizing the modified cost function can be prohibitive. Bocquet and Sakov [64] showed that when $h(x)$ is linear, the minimizer can be easily found by splitting the optimization problem in two: a scalar non-quadratic dual cost function on a compact interval, and a Kalman-like quadratic function.

Due to these challenges, an ad-hoc correction to under-sampling has become common. The technique is a combination of *covariance localization* (CL) and *covariance inflation* (CI). The issues raised by the spurious cross-correlation is usually observed when estimating dynamical systems in which the variables become spatially uncorrelated beyond some characteristic separation (e.g., a length scale in turbulent flows). It is statistically expected that small ensemble sizes will not be able to accurately capture this vanishing cross-correlation, and degrade the performance of the estimator. Hamill, Whitaker, and Snyder [66] proposed a localization scheme that explicitly enforces a decorrelation length in the computed cross-covariance matrix:

$$\hat{C}_k = L \cdot (\hat{A}_k \hat{A}_k^T), \tag{1.66}$$

where

$$L_{i,j} = \exp \left(-\frac{d_{ij}^2}{l_{corr}^2} \right) \tag{1.67}$$

is a localization matrix whose entries enforces a decorrelation pattern onto the ensemble-evaluated cross-covariance matrix. The dot product here corresponds to the element-wise multiplication, $d_{i,j}$ corresponds to the spatial distance between the ith and jth entries in the state vector, and l_{corr} is some characteristic decorrelation length for the problem in scrutiny.

The covariance inflation, on the other hand, artificially increases the ensemble covariance in order to weight the measurement data more heavily. Kelly, Law, and Stuart [67] showed that for a large enough inflation, the boundedness of the EnKF can be guaranteed. In general, the covariance inflation can be implemented as

$$\hat{x}^{(j)} = \bar{\hat{x}} + \alpha(\hat{x}^{(j)} - \bar{\hat{x}}) + \beta^{(j)} , \qquad (1.68)$$

where $\beta^{(j)}$ is the additive covariance inflation vector that is usually drawn from a zero-mean normal distribution with covariance S, and α is the multiplicative covariance inflation parameter. Both Whitaker and Hamill [68] and Bocquet, Raanes, and Hannart [65] suggested that multiplicative inflation is especially useful in mitigating the pernicious effects of the sampling errors associated with a small ensemble while additive inflation seems to be most effective in capturing sources of error that do not depend on the assimilation process such as the system modeling error. However, any practical implementation of an additive scheme requires prior knowledge of this modeling error (namely, a way of producing the matrix S). Multiplicative CI can delay the collapse of the covariance, while using additive CI will enforce a lower bound to the system covariance. In addition, the inflation parameter must be chosen carefully to avoid a divergent result.

The simplest multiplicative covariance inflation scheme is the one suggested by Anderson and Anderson [54] (AA), in which α is a scalar (typically, $\alpha \in [1.005, 1.05]$). This scheme corresponds to use $g_k(x_k) = \sqrt{\hat{C}_k}$. Its effect on the analysis scheme can be better understood by considering a slightly modified version of Eq. 1.51 that now includes the inflation parameter α. Using the posterior mode as a proxy for the posterior mean, we seek to minimize

$$\begin{aligned} J(x) &= \frac{1}{2} \|y_k - h(x)\|_R^2 + \frac{1}{2\alpha} \|x - \hat{x}_k\|_{\hat{C}_k}^2 \\ &= \frac{1}{2}[y_k - h(x)]^T R^{-1}[y_k - h(x)] + \frac{1}{2\alpha}[x - \hat{x}_k]^T \hat{C}_k^{-1}[x - \hat{x}_k] , \qquad (1.69) \end{aligned}$$

where the first term penalizes the data mismatch between the observed measurement y_k and the ones predicted by the proposed observation model, and the second term

penalizes the distance from the estimate forecast by the proposed dynamical model. The relative importance between these two models is prescribed by the matrices R (measurement noise variance) and $\hat{C}_k$ (ensemble prior variance) that represent, respectively, the level of reliability that is attributed to the data and the forecast model. Note that the multiplicative covariance inflation acts by reducing the relative importance of the second term in the cost function. As the filter is led to have a decrease sense of reliability in the prior, corrections tends to be more aggressive when inflation is present. Also, according to Luo and Hoteit [69], there is a direct connection between the multiplicative covariance inflation and robustness of the solution (in the H_∞ sense).

After analyzing the effect of sampling errors introduced by limited-size ensembles, Sacher and Bartello [70] concluded that more inflation is needed when observations lead to large corrections to the estimate. Later, Whitaker and Hamill [68] proposed an inflation scheme termed *relaxation-to-prior spread* (RTPS). In this case, α takes the form of the vector

$$\alpha_i = 1 + \theta \left(\frac{\sigma_i^b - \sigma_i^a}{\sigma_i^a} \right) , \tag{1.70}$$

where θ is a scalar (typically, $\theta \in [0.5, 0.95]$), and σ_i^b and σ_i^a are, respectively, the prior and posterior ensemble standard deviation for the i-th state variable. Note that, because α is now a vector, its multiplication with the perturbation vector $(\hat{x}^{(j)} - \bar{x})$ must be performed component-wise.

1.5.5 Enforcing Constraints

The corrections issued by the classical EnKF to each of the ensemble members are drawn from the subspace spanned by the perturbation matrix $\hat{A}_k$. Whatever linear constraints enforced to the prior ensemble members will, therefore, be preserved on the posterior ensemble. That is the case for the divergence-free and no-slip conditions in the fluid flow problems considered in this thesis.

This property, however, can be disrupted by ad-hoc covariance inflation and localization schemes. Since the AA CI scheme results in particles that are linear combinations of the previous ones, they will naturally satisfy any kinematic constraints (boundary conditions). That is not the case with the RTPS scheme, which may yield non-conforming particles since each state variable is updated independently. We present an *a posteriori* fix for this error in the context of fluid flows in section 2.1.

1.5.6 Nonlinear Observation Function

As shown in section 1.5.1, the usual EnKF algorithm is suitable for handling both linear and nonlinear dynamical models. In both cases, the forecast step is carried out by applying the dynamical model to each of the ensemble members. When non-linearities are present, however, there is no guarantees of preserving the Gaussianity of the state. In general, an error is introduced here by only tracking the first two moments of the underlying PDF.

In the analysis step, the formal goal is to find the mean of the posterior distribution, as this is the state that minimizes the mean square error. Since the posterior mean is cumbersome to evaluate, the common approach is to approximate it by the posterior mode. Thus, we must search for the minimizer of Eq. 1.69 that belongs to the affine subset generated by the prior estimate $\hat{x}_k$ and ensemble perturbation matrix $\hat{A}_k$.

The solution to this optimization problem must fulfill the zero-gradient criterium

$$DJ(x) = -\left[\frac{\partial h}{\partial x}(x)\right]^T R^{-1}[y_k - h(x)] + \frac{1}{\alpha}\hat{C}_k^{-1}[x - \hat{x}_k] = 0 . \qquad (1.71)$$

When the observation function is linear, i.e., $h(x) = H(x)$, $J(x)$ is a quadratic in x and the $DJ(x)$ has a single root given by

$$x_k = \underset{x \in \hat{x}_k + \text{span}(\hat{A}_k)}{\arg\min} \; J(x) \qquad (1.72a)$$

$$= \hat{x}_k + \left[H^T R^{-1} H + (\alpha \hat{C}_k)^{-1})\right]^{-1} H^T R^{-1}(y_k - H\hat{x}_k) \qquad (1.72b)$$

$$= \hat{x}_k + \alpha \hat{C}_k H^T \left[R + \alpha H \hat{C}_k H^T\right]^{-1} (y_k - H\hat{x}_k) . \qquad (1.72c)$$

However, when $h(x)$ is nonlinear, $J(x)$ is no longer quadratic and may not be convex or have a single minimum. Furthermore, as the gradient $H(x) = \frac{\partial h}{\partial x}(x)$ is now state dependent, Eq. 1.72 cannot be used to directly compute the minimizer of the cost function. Several remedies have been proposed, which are discussed next.

Implicit Linearization

This approach was proposed in an appendix of Evensen [19]. We start by creating an augmented state vector that merges both the state and the predicted measurements. The new observation function simply selects the second part of the state vector and

is, therefore, linear.

$$z_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix} = \tilde{f}(z_{k-1}) + \tilde{\mu}_k$$

$$= \begin{bmatrix} f(x_{k-1}) \\ h(f(x_{k-1}) + \mu_k) \end{bmatrix} + \begin{bmatrix} \mu_k \\ \nu_k \end{bmatrix} \tag{1.73}$$

$$y_k = \begin{bmatrix} 0 & I \end{bmatrix} z_k = G z_k . \tag{1.74}$$

The associated cost function is

$$J(z) = \frac{1}{2} \| z - \hat{z}_k \|^2_{\hat{C}_k^{zz}} + \frac{1}{2\alpha} \| y_k - G z \|^2_R , \tag{1.75}$$

where

$$\hat{C}_k^{zz} = \begin{bmatrix} \hat{C}_k^{xx} & \hat{C}_k^{xy} \\ (\hat{C}_k^{xy})^T & \hat{C}_k^{yy} \end{bmatrix} \tag{1.76a}$$

$$= \frac{1}{q-1} \sum_{j=1}^{q} (\hat{z}_k^j - \bar{z}_k)(\hat{z}_k^j - \bar{z}_k)^T \tag{1.76b}$$

$$= \hat{A}_k^z \left(\hat{A}_k^z \right)^T . \tag{1.76c}$$

Since this function is quadratic in z, the minimizer is given by

$$z_k = \underset{z \in \hat{z}_k + \mathrm{span}(\hat{A}_k^z)}{\arg\min} J(z) = \hat{z}_k + \left[G^T R^{-1} G + (\alpha \hat{C}_k^{zz})^{-1}) \right]^{-1} G^T R^{-1} (y_k - G\hat{z}_k) \tag{1.77}$$

$$= \hat{z}_k + \alpha \hat{C}_k^{zy} \left[R + \alpha \hat{C}_k^{yy} \right]^{-1} (y_k - h(\hat{x}_k)) . \tag{1.78}$$

Projecting this solution back to the state space:

$$x_k = \hat{x}_k + \alpha \hat{C}_k^{xy} \left[R + \alpha \hat{C}_k^{yy} \right]^{-1} (y_k - h(\hat{x}_k)) . \tag{1.79}$$

It can be shown that this approach corresponds to an implicit linearization about the ensemble mean (see Apendix A). Therefore the error associated with this approximation scales with $\hat{C}_k$, as the variance is a measure of the distance between the particles and the ensemble mean, as well as $h_{xx}(\bar{x})$, which measures how strong the nonlinearities are. This approach should work well as long as $h(x)$ is a monotonic function of the state (at least locally around the ensemble mean) and is not strongly nonlinear. The residual $\|Gz_k - h(x_k)\|_2$, i.e., the difference between the analyzed measurement to the observation operator applyied to the analized state, is a measure of the approximation introduced by this algorithm, as this quantity is expected to be zero when linear observation functions are employed [19].

Iterative Optimization

Here our approach is to directly minimizing Eq. 1.69 in order to find the posterior mode. Several schemes are proposed in the literature to solve this problem. In the following sections, we present these approaches as well as an improved methodology.

Newton-Raphson GMRES solver The Newton-Raphson (NR) method is a root-finding algorithm that is can be employed to search for roots of Eq. 1.71. Starting with a initial guess x^0, successive estimates of the root are obtained by the recurrence formula

$$D^2 J(x^i) \Delta v^i = -DJ(x^i)$$
$$x^{i+1} = x^i + \Delta x^i \ . \tag{1.80}$$

Evaluating the NR increment requires the computation and inversion of the Hessian of the cost function. The exact evaluation of $D^2 J$, which requires the evaluation of $h_{xx}(x)$, is often nontrivial and practically unfeasible. Instead, we use Eq. 1.52 to rewrite the cost function as

$$J(v) = \frac{1}{2} \|v\|^2 + \frac{1}{2} \left\| y_k - h\left((\hat{x}_k^{(j)} + \hat{A}_k v)\right) - v_k^{(j)} \right\|_R^2 \tag{1.81}$$

and its first derivative is given by

$$DJ(v) = v - \left[\frac{\partial h}{\partial x}\left(\hat{z}_k^{(j)} + \hat{A}_k v\right) \hat{A}_k \right]^T R^{-1} \left(y_k - h\left(\hat{z}_k^{(j)} + \hat{A}_k v\right) - v_k^{(j)}\right) \ , \tag{1.82}$$

which only requires that we know the result of applying $\partial h(x)/\partial z$ to the columns of $\hat{A}_k$, something that is usually available in the form of a linearized observation model. Then, following the example of Ahuja and Rowley [2], the second derivative $D^2 J$ can be obtained using the finite difference formula

$$D^2 J(v^*) \Delta x = \frac{DJ(v^* + \epsilon \Delta v) - DJ(v^* - \epsilon \Delta v)}{2\epsilon} \tag{1.83}$$

for a sufficiently small ϵ.

The GMRES scheme can then be used to iteratively solve for the NR increment Δv^i. Because this approach requires two layers of nested iterative loops (GMRES + NR), the analysis step becomes prohibitively expensive.

Gauss-Newton solver Since the cost function (Eq. 1.69) we want to minimize is a sum of squares, we can avoid the need of computing the second derivatives of $h(x)$ by employing Gauss' approximation to the Hessian matrix

$$D^2 J(v) \approx I + \left[\frac{\partial h}{\partial x} \left(\hat{z}_k^{(j)} + \hat{A}_k v \right) \hat{A}_k \right]^T R^{-1} \left[\frac{\partial h}{\partial x} \left(\hat{z}_k^{(j)} + \hat{A}_k v \right) \hat{A}_k \right] . \qquad (1.84)$$

Substituting this expression back into the NR formula, one obtains

$$
\begin{aligned}
v^{i+1} &= v^i - \left(I + B(v^i)^T R^{-1} B(v^i) \right)^{-1} \left[v^i - B(v^i)^T R^{-1} \left(y_k - h \left(\hat{z}_k^{(j)} + \hat{A}_k v^i \right) - v_k^{(j)} \right) \right] \\
&= \left(I + B(v^i)^T R^{-1} B(v^i) \right)^{-1} \times \\
&\qquad \left[\left(I + B(v^i)^T R^{-1} B(v^i) \right) v^i - v^i + B(v^i)^T R^{-1} \left(y_k - h \left(\hat{z}_k^{(j)} + \hat{A}_k v^i \right) - v_k^{(j)} \right) \right] \\
&= \left(I + B(v^i)^T R^{-1} B(v^i) \right)^{-1} B(v^i)^T R^{-1} \left[y_k - h \left(\hat{z}_k^{(j)} + \hat{A}_k v^i \right) + B(v^i)v^i - v_k^{(j)} \right]
\end{aligned}
$$

$$(1.85a)$$

$$= B(v^i)^T \left(R + B(v^i) B(v^i)^T \right)^{-1} \left[y_k - h \left(\hat{z}_k^{(j)} + \hat{A}_k v^i \right) + B(v^i)v^i - v_k^{(j)} \right] , $$

$$(1.85b)$$

where

$$B(v^i) = \left[\frac{\partial h}{\partial x} (\hat{x}_k + \hat{A}_k v^i) \hat{A}_k \right] . \qquad (1.86)$$

This iteration is very similar to Iterated Extended Kalman Filter (IEKF) scheme [36, 71], but restricted to the ensemble span. Due to this restriction, an explicit representation of $\partial h(x)/\partial x$ (or its adjoint) is not required. The evaluation of $B(v)$ only requires that we know the result of applying $\partial h(x)/\partial x$ to the columns of $\hat{A}_k$. If even this is not available (in the form of a linearized model, for instance), one can still estimate it using Eq. 1.83.

This algorithm converges in one iteration when $h(x)$ is linear, but has no convergence guarantee in the nonlinear case. Bjorck [72] showed, however, that the increment Δv obtained by the Gauss-Newton scheme is always a descent direction for J, and, if the algorithm converges, it will converge to a local extremum of J. Therefore, a damped increment $\beta \Delta v$, where $0 < \beta < 1$ is determined using a line search algorithm, can

be used to ensure convergence.

$$
\begin{aligned}
v^{i+1} &= (1 - \beta)v^i + \beta \left(I + B(v^i)^T R^{-1} B(v^i) \right)^{-1} B(v^i)^T R^{-1} \\
&\qquad \left[y_k - h \left(\hat{z}_k^{(j)} + \hat{A}_k v^i \right) + B(v^i)v^i - v_k^{(j)} \right] \qquad (1.87a) \\
&= (1 - \beta)v^i + \beta B(v^i)^T \left(R + B(v^i)B(v^i)^T \right)^{-1} \\
&\qquad \left[y_k - h \left(\hat{z}_k^{(j)} + \hat{A}_k v^i \right) + B(v^i)v^i - v_k^{(j)} \right] , \qquad (1.87b)
\end{aligned}
$$

where β is the first value in the sequence $1, 1/2, 1/4, \dots, 2^{-n}$ for which the Armijo-Goldstein step principle[72]

$$
\|J(v^{i+1})\|^2 \le \|J(v^i)\|^2 - \frac{1}{2}\beta \left(\|\Delta v^i\|^2 + \|R^{-1/2}B(v^i)\Delta v^i\|^2 \right) \qquad (1.88)
$$

is satisfied.

The iterative process ends when $\|v^{i+1} - v^i\|_\infty \le \epsilon_1$, $J(v^i) - J(v^{i+1}) \le \epsilon_2 J(v^i)$ or the pre-set maximum number of iterations is exceeded I_{max}. In this work, $\epsilon_1 = 10^{-6}$, $\epsilon_2 = 10^{-4}$ and $I_{max} = 10$.

Dealing with nonlinear observation functions in the context of Kalman filtering by using iterative schemes is not a new endeavor. The iterated Kalman filter (IKF) proposed by Jazwinski [36] follows a very similar algorithm, and was later interpreted as a Gauss-Newton scheme [73] and a Picard iteration [74]. Zupanski [75] was possibly the first researcher to propose a iterative scheme in the context of ensemble-based estimators. He proposed a variant of the ensemble transform Kalman filter (ETKF[76]) to minimize a cost function (or maximize the corresponding likelihood function) similar to Eq. 5.40. Gu, Oliver, et al. [77] later suggests an iterative Gauss-Newton update formula for the EnKF in which the observation function was linearized about each of the intermediate ensemble means.

The major differences between the algorithm proposed here and the one of Gu, Oliver, et al. [77] are:

- Linearization is performed about the current estimate of the conditional mode for each of the ensemble members as opposed to using the ensemble mean. As a consequence, convergence is usually obtained with 5 iterations or fewer (as opposed to twice that, if the linearization about the ensemble mean is used)

- Since ensemble statistics no longer needs to be evaluated after each iterative step in order to update the linearization, the analysis step can be performed in parallel

Another important detail is that when nonlinearities are present, there is no guarantee that the posterior remains a Gaussian and, therefore, the mode(s) and the mean may no longer coincide. The magnitude of this discrepancy, which may be considered yet another bias source, scales with the error variance, and therefore is expected to decrease as the estimator converges.

1.6   Outline of the Contributions in this Thesis

In this thesis we propose to apply the Ensemble Kalman Filter to the problem of estimating both the flowfield and the forces exerted on a body by a low-Re flow from measurements that would be available in an experimental setting. In Chapter 2, we describe the numerical solver that will be used as both the estimator model and the surrogate for the truth. In addition to a brief summary of the Immersed Boundary Lattice Green Function (IBLGF) method that is used to simulate the flow, we show how this methodology can be extended to obtain the relations that are required for integrating a IBLGF-based model into the EnKF. The chapter then concludes by presenting an overview of the numerical setup of the different flows that will be considered throughout the thesis.

In Chapter 3, we survey the characteristics of the EnKF estimator under idealized conditions, i.e. in the absence of modeling errors. We start by accessing the impact of measurement noise levels, and schemes for initializing the ensemble, performing covariance inflation, and dealing with nonlinear observation functions. Using Evensen's representers formulation, we recall that the filter corrections are related to the ensemble cross-covariance matrix between the state and the measurements. This matrix is shown to reproduce some of the conclusions of structural sensitivity analysis for the same types of flows, and is used to propose a methodology for optimal sensor placement. The optimality of the implemented estimator is accessed for both pressure and velocity measurements by analyzing the whiteness of the respective innovation sequence. The Chapter concludes with a comparison between EnKF and 3D-Var methodologies.

Since the computational cost of forecast and observation models have a significant impact on the runtime of these estimators, dealing with modeling errors is a practical

inevitability. When these errors have non-zero mean and are left unaccounted for, the introduced bias can severely impair the estimator performance. In Chapter 4 we investigate one source of modeling error, namely uncertainty (or misspecificiation) of model parameters. Two strategies to deal with these errors are discussed: agnostic and state augmentation approaches. In the first test case we emulate misspecified boundary condition (BC), represented by different kinds of freestream perturbations, and the challenge is to be able to continuously[12] estimate the freestream velocity using only pressure measurements on the surface of the airfoil. In the second test case we emulate a misspecified initial condition (IC), and use use the state augmentation approach to estimate the Reynolds number in the reference run. Our ability to estimate a critical parameter on the fly shows how the EnKF can overcome a common fragility displayed by estimators based on reduced-order models.

Sometimes, however, model errors must be dealt with by treating their effects (mitigating the resulting bias) rather than their sources (fixing the model). In Chapter 5, we examine a more challenging modeling error, namely the use of under-resolved models[13] to estimate real flows. A direct representation of these errors can be especially challenging for ensemble methods. In order to overcome these limitations, we propose a low-rank representation for the modeling error, and use colored-noise processes to represent the dynamics of the slow-varying portion of the bias. The Ensemble Kalman Filter framework is then employed to simultaneously correct both the state and bias parameters. The proposed methodology is demonstrated using the twin-experiment strategy: using a Navier-Stokes equations solver with immersed boundaries capabilities, the state of a fine-grid two-dimensional low-Re flow simulation past an inclined flat plate is estimated using an ensemble of coarse-mesh simulations and pressure measurements taken on the surface of the plate. Finally, Chapter 6 summarizes the most important findings and proposes some avenues for future work.

---

[12]Differently from the misspecified IC case, the estimator needs to remain responsive to further changes in the freestream at later estimation times.

[13]Even though one could minimize this error by discretization refinement, the associated additional cost is undesirable and often prohibitive.

C h a p t e r   2

NUMERICAL METHOD AND FLOWS CONSIDERED

2.1   The Numerical Method

The dynamics of the flow are represented by the discretized 2-D incompressible Navier-Stokes equations, and the simulations were carried out using the Immersed Boundary Projection Method (IBPM)[78, 79] enhanced by the Lattice Green's Function (LGF) formulation [80, 81]. The LGF formulation exactly enforces the free-space boundary condition at infinity for the Poisson equation even though the computation domain is restricted to a snug region of non-zero vorticity near the immersed body.

2.1.1   The Immersed Boundary Projection Method (IBPM) in an Accelerating Frame

Consider the non-dimensional incompressible Navier-Stokes equations written in the inertial frame with a singular immersed boundary force $\mathbf{f}$ added to the momentum equation

$$\frac{\partial \mathbf{u}_f}{\partial t} + \mathbf{u}_f \cdot \nabla \mathbf{u}_f = -\nabla \cdot p + \frac{1}{Re} \nabla^2 \mathbf{u}_f + \int_\Gamma \mathbf{f}(\xi(s,t))\delta(\xi - \mathbf{x_f})ds \tag{2.1a}$$

$$\nabla \cdot \mathbf{u}_f = 0 \tag{2.1b}$$

$$\mathbf{u}_f(\xi(s,t)) = \int_\mathbf{x} \mathbf{u}_f(\mathbf{x})\delta(\mathbf{x} - \xi)d\mathbf{x} = \mathbf{u}_b(\xi(s,t)) \,, \tag{2.1c}$$

where $\mathbf{u}$ and p are the velocity and pressure variables. Now, consider a accelerating frame that moves with linear velocity $\mathbf{U}_f(t)$, and rotates about the origin of accelerating frame with angular velocity $\mathbf{\Omega}(t)$, as described by Tsai and Colonius [82]. The relationship between the velocities and vorticity evaluated in each of the two reference frames is given by

$$\mathbf{u}_f = \mathbf{u}_r + \mathbf{\Omega}(t) \times \mathbf{x}_r + \mathbf{U}_f(t) \tag{2.2}$$

$$\omega_f = \omega_r + 2\mathbf{\Omega}(t) \,, \tag{2.3}$$

where the subscripts f and r denote variables in the inertial and accelerating frames, respectively.

Therefore, the momentum equation written in the accelerating frame is given by [82]

$$\left(\frac{\partial \mathbf{u}_r}{\partial t}\right)_r + \mathbf{u}_r \cdot \nabla \mathbf{u}_r = -\nabla \cdot p + \frac{1}{Re}\nabla^2 \mathbf{u}_r$$

$$- \dot{\mathbf{\Omega}}(t) \times \mathbf{x}_r - 2\mathbf{\Omega}(t) \times \mathbf{u}_r - \mathbf{\Omega}(t) \times (\mathbf{\Omega}(t) \times \mathbf{x}_r) - \left(\frac{\partial \mathbf{U}_f}{\partial t}\right)_f$$

$$+ \int_\Gamma \mathbf{f}(\xi(s,t))\delta(\xi - \mathbf{x_r})ds \ . \tag{2.4}$$

The boundary conditions for Eq. 2.4 are [83]

$$\begin{cases} \mathbf{u}_r(t) \rightarrow -\mathbf{U}_f(t) - \mathbf{\Omega}(t) \times \mathbf{x}_r \\ \omega_r(t) \rightarrow -2\mathbf{\Omega}(t) \qquad\qquad\quad \text{as } \|\mathbf{x}_r\| \rightarrow \infty \\ p(t) \rightarrow p_\infty(t) \end{cases}$$

$$\mathbf{u}_r(t) = 0 \qquad\qquad\qquad\qquad \text{on } \Gamma \ .$$

In order to make the boundary conditions consistent with the LGF framework, we use the change of variables proposed by Tsai [83]:

$$\mathbf{u}' = \mathbf{u}_r + \mathbf{\Omega}(t) \times \mathbf{x}_r + \mathbf{U}_f(t) \tag{2.5a}$$

$$\omega' = \omega_r + 2\mathbf{\Omega}(t) \tag{2.5b}$$

$$p'_t = p_t - p_\infty - \frac{1}{2}\|\mathbf{U}_f(t)\|^2 \ , \tag{2.5c}$$

where

$$p_t = p + \frac{1}{2}\|\mathbf{u}_r\|^2 - \frac{1}{2}\|\mathbf{\Omega}(t) \times \mathbf{x}_r\|^2 - \mathbf{U}_f \cdot (\mathbf{\Omega}(t) \times \mathbf{x}_r) \tag{2.6}$$

is the total pressure.

By taking the curl of Eq. 2.4 and applying the change of variables, we obtain the vorticity equation

$$\frac{\partial \omega'}{\partial t} = \nabla \times (\mathbf{u}_r \times \omega') - \frac{1}{Re}\nabla \times (\nabla \times \omega') + \nabla \times \left(\int_\Gamma \mathbf{f}(\xi(s,t))\delta(\xi - \mathbf{x_r})ds\right) \ . \tag{2.7}$$

Similarly, by taking the divergence of Eq. 2.4 and applying the change of variables, we obtain the pressure equation

$$\nabla^2 p'_t = \nabla \cdot (\mathbf{u} \times \omega') + \nabla \cdot \left(\int_\Gamma \mathbf{f}(\xi(s,t))\delta(\xi - \mathbf{x_r})ds\right) \ , \tag{2.8}$$

subjected to the boundary conditions

$$
\begin{cases}
\mathbf{u}'(t) \to 0 \\
\omega'(t) \to 0 \qquad\qquad \text{as } \|\mathbf{x}_r\| \to \infty \\
p'_t(t) \to 0
\end{cases}
$$

$$
\mathbf{u}'(t) = \mathbf{\Omega}(t) \times \mathbf{x}_r + \mathbf{U}_f(t) \quad \text{on } \Gamma \,.
$$

These equations are discretized using a staggered-mesh finite-volume formulation [78]. In this framework, different flow quantities are evaluated on distinct positions within the unit cell. Figure 2.1 shows how the velocity components u and v, the pressure p, the vorticity ω and the streamfunction ψ are placed with respect to the primal cell and its respective dual. We will consider a uniform grid with equal grid spacing h in the x and y directions.
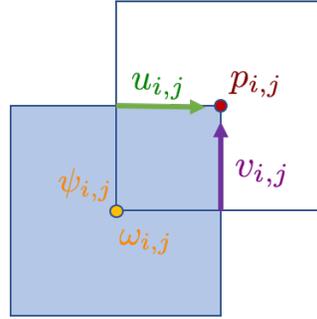
Figure 2.1: Variables placement on a staggered grid with respective dual mesh.

By appropriately discretizing the continuous differential operators in Eq. 2.7, the following semi-discrete DAE system is obtained

$$
\frac{\partial \omega'}{\partial t} = C^T N(\omega') - \frac{1}{Re_\Delta} C^T C \omega' - C^T H f \tag{2.9a}
$$

$$
EC(C^T C)^{-1} \omega' = E\mathbf{u}' = -\mathbf{\Omega}(t) \times \mathbf{x}_r - \mathbf{U}_f(t) \,, \tag{2.9b}
$$

where C is the discrete curl operator, $H = \frac{\Delta s}{h} E^T$ is the regularization operator, E is the interpolation operator, $Re_\Delta = Re \frac{h}{L}$ is the grid Reynolds number and $N(\omega') = \mathbf{u}_r \times \omega'$ is the discretization of the nonlinear advection term. The surface stresses f here represent the action of the fluid on the immersed body (hence the minus sign). Details about the construction of the discrete operators can be found in the Appendix A of Colonius and Taira [79].

The matrices H and E are constructed to represent a discrete version of the convolution integral that appears in both the vorticity equation and the IB constraint. The

continuous delta function is approximated by its discrete counterpart, the discrete delta function (DDF). There are several options of DDFs in the literature (see Fig. 2.2 for some examples) that differ from each other by their support (how many cells around the IB point the force is spread out and the velocity is interpolated from) and degree of smoothness (how many continuous derivatives the DDF has). The balance between these two features will determine the smoothness of the evaluated IB forces and the perceived sharpness of the IB for a given grid resolution. In this work, we use the smoothed 3-point DDF of Yang et al. [84]

$$\phi_3^*(r) = \begin{cases} \frac{17}{48} + \frac{\sqrt{3}\pi}{108} + \frac{|r|}{4} - \frac{r^2}{4} + \frac{1-2|r|}{16}\sqrt{-12r^2 + 12|r| + 1} \\ \qquad - \frac{\sqrt{3}}{12} \arcsin\left(\frac{\sqrt{3}}{2}(2|r|-1)\right) & |r| \leq 1 \\ \frac{55}{48} - \frac{\sqrt{3}\pi}{108} - \frac{13|r|}{12} + \frac{r^2}{4} + \frac{2|r|-3}{48}\sqrt{-12r^2 + 36|r| - 23} \\ \qquad + \frac{\sqrt{3}}{36} \arcsin\left(\frac{\sqrt{3}}{2}(2|r|-3)\right) & 1 \leq |r| \leq 2 \\ 0 & |r| \geq 2 \, . \end{cases}$$

(2.10)

It was also noted by Goza et al. [85] that using a $1 < \frac{\Delta s}{h} < 2$ is beneficial with regard to producing smoother force distributions.
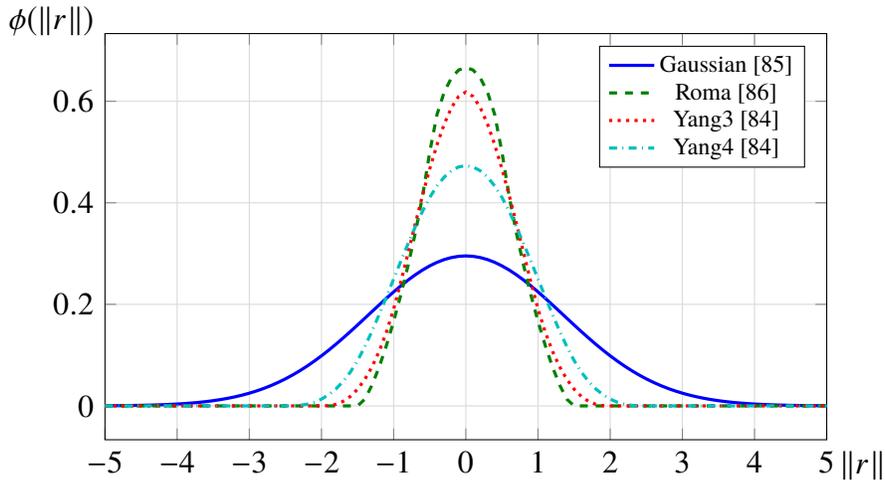
Figure 2.2: Comparison between different choices of the discrete delta function.

The evaluation of the nonlinear term requires the procurement of the velocity field from the vorticity field. This is accomplished via the evaluation of the streamfunction ψ using the lattice Green's function.

$$\omega' = C^T \mathbf{u}' = C^T C \psi \tag{2.11}$$

$$\mathbf{u}' = -CL^{-1}\omega' \, , \tag{2.12}$$

where $-C^T C$ corresponds to the scalar Laplacian operator operating over the cell vertices. The solution to this discrete Poisson equation has an explicit analytic solution in the form

$$u(\mathbf{m}) = [\phi \star f](\mathbf{m}) = \sum_{\mathbf{p} \in \mathbb{Z}^2} G(\mathbf{m} - \mathbf{p}) f(\mathbf{p}) , \qquad (2.13)$$

where m is the mesh index vector, and G is the fundamental solution of the Poisson equation in an infinite Cartesian mesh (see Appendix B for more details). Assuming that the regions with nonzero vorticity are compact, this convolution can be evaluated efficiently using the fast Fourier transform (FFT) and the convolution theorem.

Two time marching scheme options are available. The first approach uses the classical second-order Crank-Nicolson/Adam-Bashforth (CNAB2) temporal split scheme [78, 87]. The second uses a integrating factor approach to avoid the stiffness introduced by the viscous terms in the Navier-Stokes equation, and a half-explicit Runge-Kutta (HERK) scheme, specially suitable to deal with systems of ODEs with algebraic constraints, is used as the time marching scheme [88]. In both cases, the resulting fully-discrete equations are solved using block-wise LU decomposition. The results for the cylinder and airfoil (with exception of section 5.9) were obtained with the HERK scheme, while the flat plate cases were run with the CNAB2 scheme.

The same time-marching framework is also used to enforce the no-slip condition in the posterior state (see section 1.5.5 for further details on this). The non-conforming part of the state can be projected out of the solution by applying a single iteration of the selected time-marching scheme with $\Delta t = 0$.

2.1.2 CFL Constraint

Since an explicit time-marching scheme will be used as part of the dynamical model of the estimator, the stability of the solution is conditioned to a CFL constraint. The corrections applied by the estimator can create regions of the solution with high velocities, which can lead to the divergence of the solution. This is the most common cause of estimator crash. Because the corrections are scaled by the ensemble variance, and are proportional to the data mismatch, this issue is more likely to occur early on in the estimation history.

There are basically two parameters that can be adjusted to prevent estimator divergence due to CFL constraint violation: the numerical time step and the covariance inflation magnitude. The first one is an obvious choice, but a smaller time step

will necessarily increase the cost of the forecast step, and therefore it is interesting to use the largest possible time step. In fact, one can use a varying time step that increases as the the data mismatch and ensemble variance decrease, saturating at a safe maximum limit. In the simulations carried out in the present work we used

$$\Delta t = \min \left(\Delta t_{max}, \frac{\|C_0\|}{\|\hat{C}_k\|} \Delta t_0 \right) , \qquad (2.14)$$

where C_0 is the initial ensemble covariance matrix, Δt_0 is a sufficiently small time step, and Δt_{max} is the time step used to run the reference simulations.

An alternative is to make corrections smaller by weighting more the a priori prediction. This can be done either by increasing the measurement noise level or decreasing the ensemble variance. To that end, it may be advisable to pull back the magnitude of covariance inflation in use.

2.1.3 The Pressure Solver

The IBLGF framework is also used to evaluate the corresponding free-space pressure field. Discretizing 2.8, one obtains

$$DGp'_t = DN(u) - DHf , \qquad (2.15)$$

where G is the discrete gradient operator (which is the dual of the divergence operator $D = G^T$).

It can be shown that DG also corresponds to the scalar Laplacian operator, but operating over the cell centers. Therefore, Eq. 2.15 correspond to a discrete Poisson equation on an infinite grid, which can be solved using the LGF. With the total pressure in hands, the pressure coefficient can be evaluated from Eq. 2.6:

$$c_p = \frac{p - p_\infty}{0.5\rho\|U_f\|^2} = 1 - \frac{\|u_r\|^2}{\|U_f\|^2} + 2\frac{p'_t}{\|U_f\|^2} + \frac{\|\Omega \times x_r\|^2}{\|U_f\|^2} + 2\frac{U_f \cdot (\Omega \times x_r)}{\|U_f\|^2} . \quad (2.16)$$

Figure 2.3 shows an example of the evaluated pressure field for the flow past a cylinder. As we would expect from a IB method, when closed bodies are present, there is also a non-unique pressure distribution corresponding to the region inside the body.

2.1.4 The Semi-Discrete Force Update

In the IBLGF framework, f plays the role of a Lagrange multiplier that enforces the no-slip boundary condition at each of the body points, and whose value is obtained
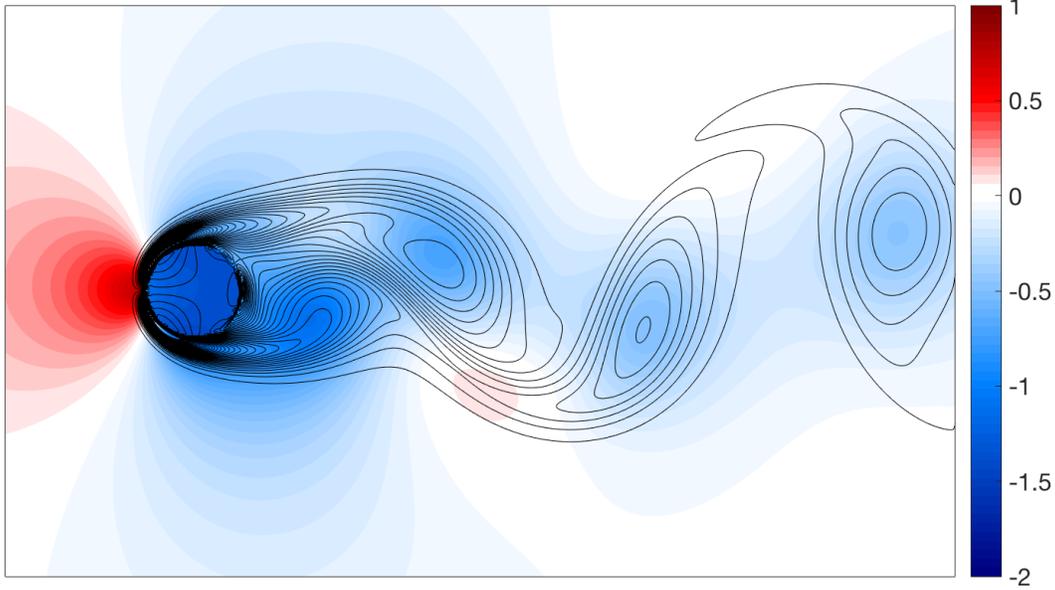
Figure 2.3: Contour plot of the computed pressure coefficient field, overlapped with isocurves of vorticity.

as a byproduct of the time-marching procedure. However, sometimes[1] it is necessary to obtain these forces from the vorticity field at a instantaneous moment in time. For that end, a semi-discrete approach is used.

Thus, by taking the time derivative of the algebraic constant equation and substituting the vorticity equation in

$$EC(C^TC)^{-1}\frac{d\omega}{dt} = \frac{du_b}{dt}$$

$$EC(C^TC)^{-1}\left(-\frac{1}{Re}C^TC\omega + C^TN(w) - C^TE^T\tilde{f}\right) = \frac{du_b}{dt}$$

$$ECL^{-1}C^TE^T\tilde{f} = ECL^{-1}\left(\frac{1}{Re}L\omega + C^TN(w)\right) + \frac{du_b}{dt}$$

$$\tilde{f}(\omega) = (ECL^{-1}C^TE^T)^{-1}\left[ECL^{-1}\left(\frac{1}{Re}L\omega + C^TN(w)\right) + \frac{du_b}{dt}\right]. \tag{2.17}$$

All these operators are already used by the time marching scheme, therefore the implementation is straight-forward. Note that this map between state (ω) and forces ($\tilde{f}$) is nonlinear due to the presence of the convective term in the right-hand side. Also, since the implemented time-marching scheme is second-order accurate (both the CNAB2 and HERK), the force so obtained is a second-order approximation to the force obtained using this semi-explicit scheme, as Fig. 2.4 indicates.

---

[1]In the EnKF context, an instantaneous update of the IB forces is necessary when process noise is present or when iterative schemes are used to deal with a nonlinear measurement function.
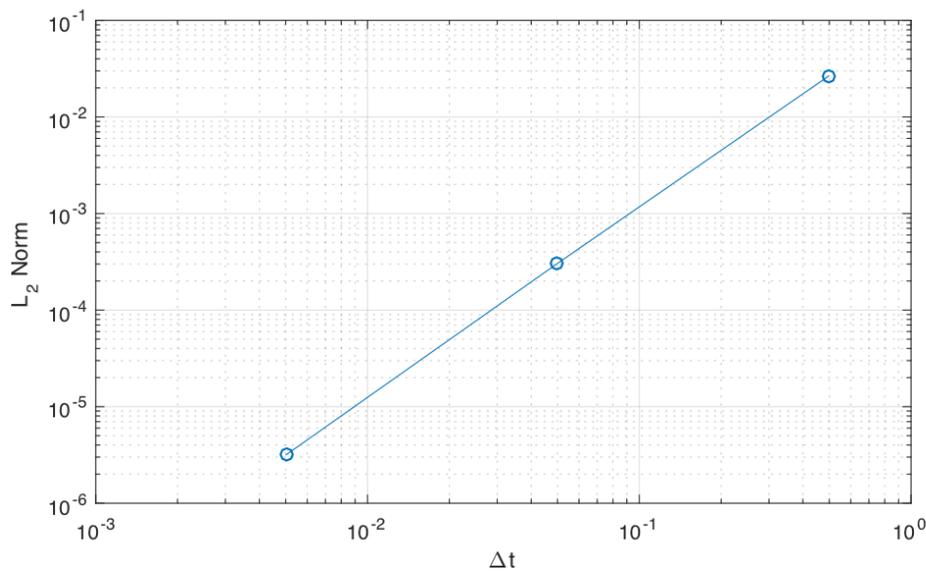
Figure 2.4: Convergence plot for the semi-discrete force update.

The forces obtained by either the time-marching or the force update are often noisy, a consequence of the ill-conditioning of the integral equation of the first kind expressed by Eq. 2.1c. Goza et al. [85] proposed a post-processing scheme that filters out the high-frequency spurious oscillations given by $f_{filt} = EWHf$, where

$$
W_{ii} = \begin{cases} 1/(H\mathbb{1})_i & (H\mathbb{1})_i \neq 0 \\ 0 & \text{otherwise .} \end{cases} \tag{2.18}
$$

Note also that the problem of finding the force acting on the IB points is ill-posed when a closed body is present. On those bodies, the IB forces are only determined up to a constant normal component (as such a force distribution would have zero resultant on the fluid). The actual value for this constant that is computed by this scheme will depend on the implicit regularization provided by the choice of discrete delta function and grid resolution. In practice, this constant normal component can be quite large and, as a consequence, the pressure inside the body (computed using the formulation presented in section 2.1.3), can achieve extreme values. This behavior can be prevented by enforcing, a posteriori, a fixed value for this constant normal component. In Chapter 3, this value was chosen to be zero[2].

---

[2]Note that individual choices have absolutely no impact in the flowfield outside the body.

2.1.5 The Linearized Force Update

Although obtaining a matrix representation of the linearized force update operator may be challenging, the IBLGF framework can be used to compute the action of the linearized operator on a vector

$$\frac{\partial \tilde{f}}{\partial \omega}(\omega_b)\omega = (ECL^{-1}C^T E^T)^{-1}\left[ECL^{-1}\left(\frac{1}{Re}L\omega + C^T\frac{\partial N}{\partial w}(\omega_b)\omega\right)\right] , \qquad (2.19)$$

where

$$\frac{\partial N}{\partial w}(\omega_b)\omega = u_b \times \omega + u \times \omega_b \qquad (2.20)$$

is the linearization of the convective term of the Navier-Stokes equation and u_b is the velocity field associated with ω_b, the solution about that the linearization takes place.

2.2 The Problem of Interest

In this work, we will be studying methodologies that are suitable of dealing with dynamical systems that share the following characteristics:

- high-dimensional numerical model

- sparse spatially-restricted observations

- nonlinear forecast and observation models

In particular, we will be interested in estimating the state of the low-Re incompressible flow past 2D canonical bodies. This flow is characterized by laminar vortex shedding with a well-defined frequency. For the cylinder at $Re = 100$, we expect a Strouhal number of $St = 0.165$[89]. For the other two geometries, the corresponding Strouhal number can be estimated using the Reynolds number associated to the projection of the chord into the direction perpendicular to the incoming flow. We set the angle of attack ($AoA = 30°$) and Reynold number ($Re = 200$) for the inclined flap plate and airfoil, so that the projected Reynolds number ($Re \sin(AoA) = 100$) is equivalent to the cylinder case. Figures 2.10, 2.9, and 2.6 exemplify these characteristics for the flat plate, NACA 0009 airfoil and the circular cylinder, respectively.

The lattice Green's function formulation enables us to restrict the computational domain to the region close to the body and the wake (nonzero vorticity region). Nevertheless, the cylinder case requires about 15000 degrees of freedom for a

grid resolution corresponding to $Re_\Delta = 4$ (which is borderline high for obtaining physically relevant results), and the flat plate and airfoil cases require 60000 degrees of freedom for the same resolution level and domain size. That notwithstanding, it is important to note that these flows exhibit low-rank dynamics. Figure 2.5 shows the fraction of the variance of the solution (computed using a sufficiently long run) that is left out by keeping only the leading POD modes. About 25 modes are enough to retain 99.9% of the variance in all cases.
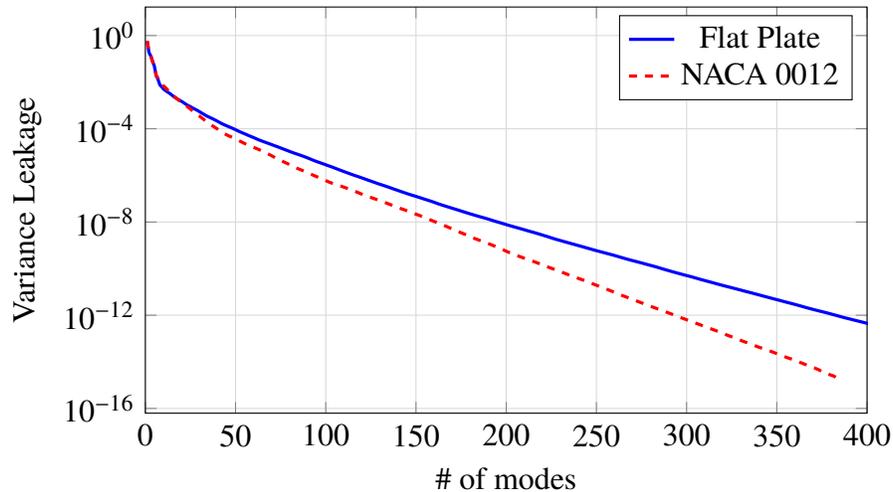
Figure 2.5: Low-rank behavior of the low-Re incompressible flow past inclined flat plate and airfoil (NACA 0012).

As for the measurements, having the future experimental applicability of the conclusions drawn by this study in mind, it is important to only observe quantities that would be accessible in a typical laboratory setting. Thus, observable quantities consist in either the velocity components sampled in several positions in the wake of the body, producing the kind of data that resemble time-resolved PIV data, or the stresses on the surface of the body, from that the pressure distribution can be estimated. While the velocity field is a linear function of the vorticity field, the force distribution acting on the immersed boundary retains a dependency on the advection term of the N-S equations, and is therefore a nonlinear function of the vorticity field. This distinction will have an important impact on the accuracy of the estimation problem.

2.2.1 Low-Re Flow Past a Circular Cylinder

The first test case is the canonical flow past a circular cylinder at $Re_D = 100$. At this Reynolds number, the flow is still essentially two-dimensional [90] and is characterized by coherent vortices in the wake known as the Karman vortex street. The dynamical system is deterministic, and, apart from synthetic noise added to the measurements, the only randomization is associated with the initial conditions. Thus, the goal is to use the EnKF to track the phase of the vortex street.
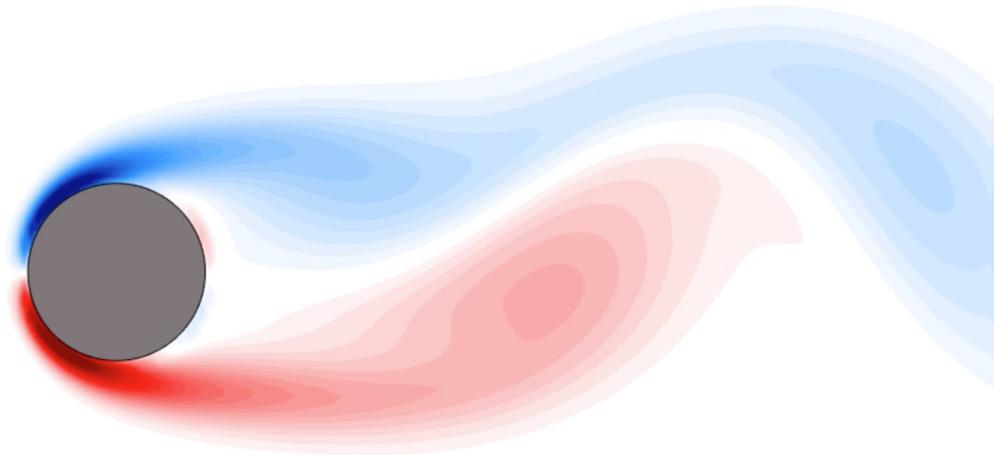
Figure 2.6: Vorticity contours for the flow past a circular cylinder at $Re = 100$.

The grid resolution is 25 points by diameter (which yields a system with about 15,000 degrees of freedom), and velocity components sampled at 14 equidistant points in the flowfield (see Fig. 2.7) are assimilated into the filtering process every 0.1 convective time units.

2.2.2 Low-Re Flow Past a NACA 0009 Airfoil

The second test case is closer to the motivation presented in the introduction section. The performance of the estimator is now analyzed when applied to the problem of the flow past a NACA 0009 at 30 degrees angle of attack and Reynolds number 200. The grid resolution is 50 points per chord and the pressure at distinct locations over
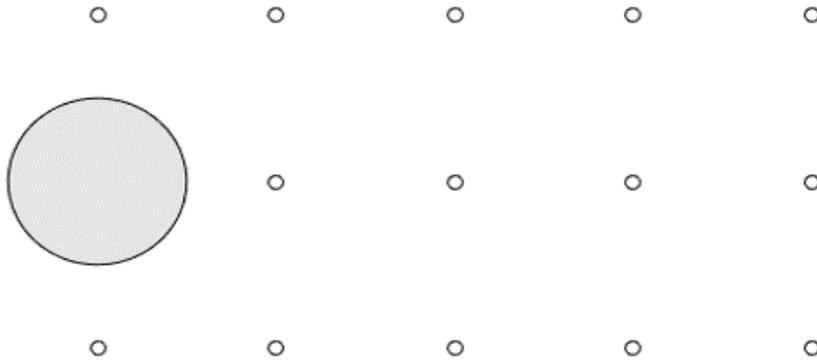
Figure 2.7: Flow past a circular cylinder: location of the velocity measurement points in the flowfield.

the airfoil (see Fig. 2.8) are taken as measurement data every 0.05 convective time units.

Figure 2.8: Location of the pressure measurement points over the surface of a NACA 0009 airfoil.

2.2.3   Low-Re Flow Past a Inclined Flat Plate

The third test case is the flow past a flat plate at 30 degrees angle of attack and Reynolds number 200. The state of this system is represented by the vorticity at each of the grid points. With the purpose of analyzing the effects of resolution, meshes with grid Reynolds numbers ($Re_\Delta = Re\Delta x/c$ where $\Delta x = \Delta y$ is the grid spacing) equal to 1 (200 points per chord), 2 (100 points per chord) and 4 (50 points per chord) are used. This simple geometry was considered in addition to the previous airfoil case to minimize changes in the body geometry as the grid resolution is refined. As for the measurements, we sense the pressure at 10 equidistant locations over the plate every 0.05 convective time units.
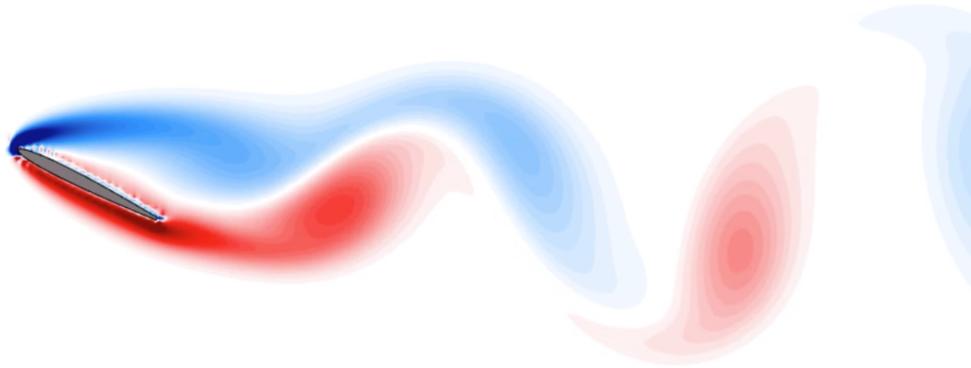
Figure 2.9: Vorticity contours for the the flow past a NACA 0009 airfoil at $AoA = 30\,\text{deg}$ and $Re = 200$.
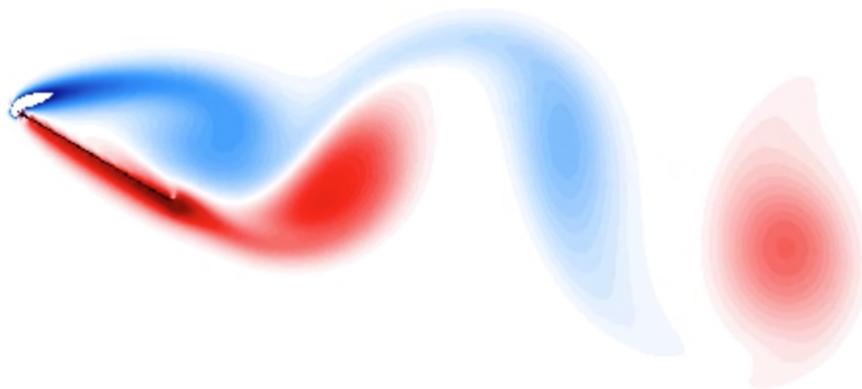
Figure 2.10: Vorticity contours for the flow past a flat plate at $AoA = 30\,\text{deg}$ and $Re = 200$.

2.3 Performance Evaluation Metrics

In order to evaluate the performance of the estimator, the following metrics will be used throughout the thesis:

- State estimate error (measures how the estimated state differs from the reference state):

$$E_x = \frac{\|\bar{x} - x_{ref}\|}{\|x_{ref}\|} \; . \tag{2.21}$$

- Observation estimate error (measures how the estimated observation differs from the measured data):

$$E_y = \frac{\|\bar{y} - y_{ref}\|}{\|y_{ref}\|} \ . \tag{2.22}$$

- State ensemble RMS (measures the spread of the ensemble state estimates):

$$RMS_x = \sqrt{\frac{1}{q-1} \sum_{i=1}^{q} \frac{\|x^{(i)} - \bar{x}\|^2}{\|\bar{x}\|^2}} \ . \tag{2.23}$$

- Observation ensemble RMS (measures the spread of the ensemble observation estimates):

$$RMS_y = \sqrt{\frac{1}{q-1} \sum_{i=1}^{q} \frac{\|y^{(i)} - \bar{y}\|^2}{\|\bar{y}\|^2}} \ . \tag{2.24}$$

Chapter 3

ESTIMATION IN A PERFECT-MODEL FRAMEWORK

The estimator performance will be significantly impacted by the choice of how to model and observe the tracked system. The definition of the predictive model must be a compromise between accuracy and cost. The predictive model should be accurate enough to alleviate the burden laid upon the error control techniques (stochastic forcing, covariance inflation, etc) while keeping the computational time controlled. On the other hand, the choice of what and where to measure will not only limit the accuracy of the state/output estimates, but will also determine if the estimation task is feasible at all (detectability[1]).

In this section, we discuss how tuning the different estimation parameters affect the performance of the estimator in a perfect model framework. We show that the representers of the estimator (defined in Eq. 1.58b) provide guidance on sensor placement. Each subsection focuses on a distinct aspect of the EnKF methodology and is sorted by the test problem that was used to perform the analyses.

3.1 Flow Past a Circular Cylinder

Throughout this section, the flow past a circular cylinder at $Re = 100$ (see section 2.2.1 for more details) was used as test problem. The grid resolution was set to 25 points by diameter ($Re_\Delta = 4$), and velocity measurements were taken at different locations in the wake every 0.1 convective time units. The AA multiplicative covariance inflation scheme (see section 1.5.4) with $\alpha = 1.05$ was used.

3.1.1 Effect of the Assimilation Interval

The frequency with which the estimator must obtain new measurements must be a compromise. A lower bound is defined by the parallel efficiency of the EnKF algorithm. Assimilation shouldn't occur too often because the analysis step is inherently a serial procedure and requires a big volume of data transfer in order to evaluate the necessary ensemble statistics.

---

[1]Detectability is a weaker notion of observability. While observability requires all nontrivial modes to produce a non-zero output, detectability only requires neutral and unstable modes to produce non-zero output.

On the other hand, the characteristics of the dynamical system impose an upper bound on the assimilation interval. New measurements must be frequent enough to ensure that the corrections issued by the estimator can overcome the divergence of the most unstable modes of the dynamical systems. It is also important to ensure that the assimilation frequency is faster than any relevant time scale of the problem.

Unless otherwise noted, the assimilation interval for the results presented in this thesis is 0.05 convective time units, which corresponds to a rate of 120 measurements per vortex shedding cycle.

3.1.2   Effects of the Initialization Scheme

Figure 3.1 shows how different choices of the initial ensemble impact the evolution of the estimate error when the ensemble size is held fixed at 24 members. In the first case, we initialize the flow field with zero-mean random numbers, while in the second case, we add random perturbations to the true mean flow. Finally, in the third case, we add the first 24 POD modes to the mean, but with coefficients randomized. In all cases, the ensemble initial variance matches the reference run variance. When the initial condition is restricted to the POD subspace, the estimator convergences within a few convective time units, whereas, unsurprisingly, the more random initial conditions take longer to converge. Physically speaking, the timescale associated with the domain size in the x-direction is about 6 units, and the estimator must wash out the random initial condition over this time period while a transient takes place that leads to different realizations of the flow with distinct vortex shedding phases.

Figure 3.2 presents the evolution of the state estimate error and ensemble RMS for different ensemble sizes. Provided that the dynamics is relatively sparse in the POD space and the most energetic subspace is captured by the initial ensemble, the transient behavior of the estimator appears to become independent of the ensemble size. This behavior indicates that the sampling error decays more rapidly than the expected $q^{-1/2}$ when a POD-based initialization is used. Another interesting behavior that is portrayed in these plots is the collapse of the variance. For smaller ensemble sizes, the ensemble RMS decreases rapidly, limiting the effectiveness of further corrections to the state as the confidence on the internal state is overestimated. As a consequence, error converges in a slower rate then when fewer ensemble members are used. However, if modeling errors were present, the increasing disregard for new external information coming from the measured data may lead the estimator to diverge.
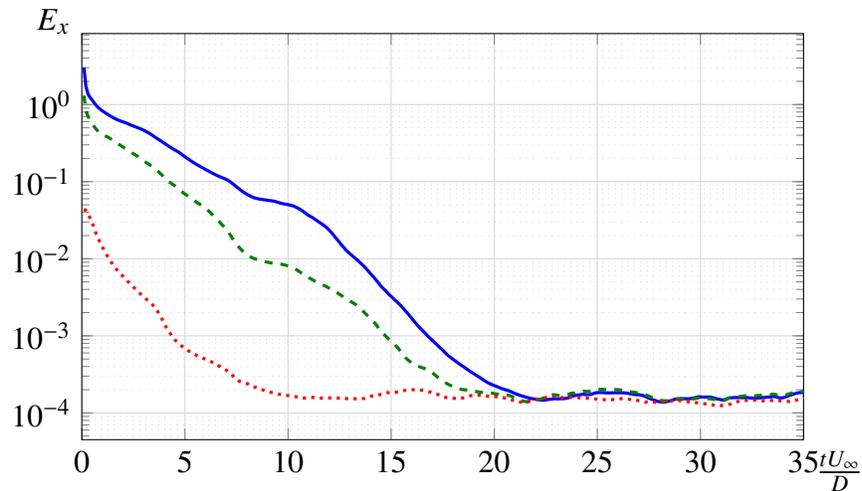
Figure 3.1: Effect of the choice of the initial ensemble on the estimator performance for the flow past a circular cylinder at $Re = 100$ using velocity measurements. The ensemble has 24 members and the measurement error level is $R = 10^{-4}I_p$. (——) corresponds to a zero-mean random initial condition, (- - -) corresponds to the mean flow plus random perturbations and ($\cdots\cdots$) corresponds to the mean flow plus randomized leading POD modes.

(a) State estimate error.

(b) Ensemble RMS.
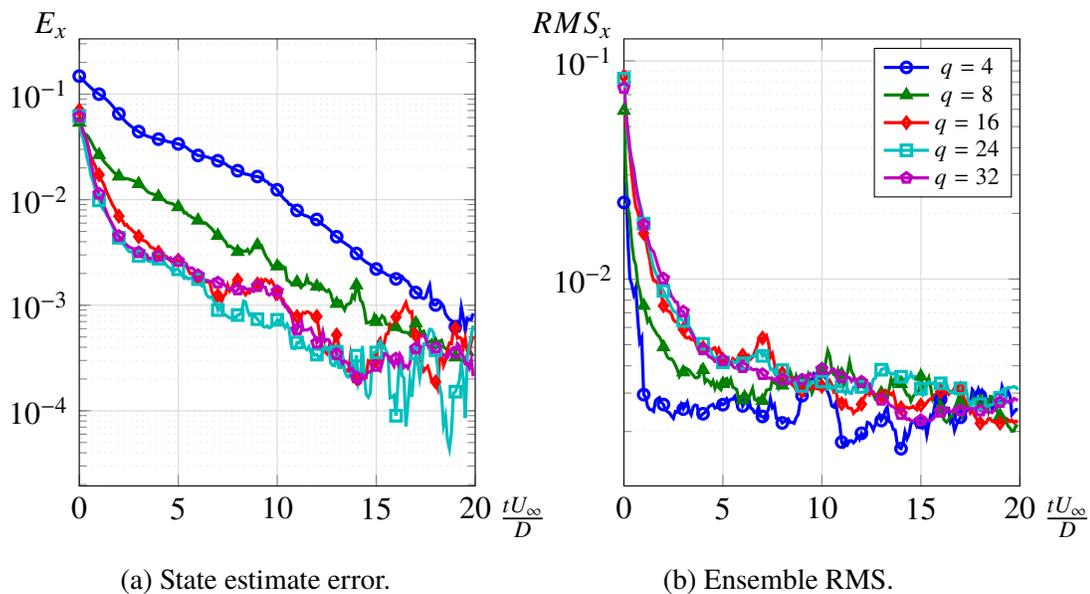
Figure 3.2: Estimator performance for increasing ensemble sizes applied to the problem of the flow past a circular cylinder at $Re = 100$ using velocity measurements. Measurement error level is set to $R = 10^{-4}I_p$.

3.1.3   Effects of the Measurement Noise Level

The measurement noise level also impacts the performance of the estimator, since the reliability of sensor data should be weighted against the internal state of the estimator. As the data assimilation proceeds, the estimated state uncertainty decreases and may reach an error level for which further corrections become secondary. Figure 3.3 shows how a lower noise level favors the estimator performance in the case of velocity sensors. Note that reducing the noise covariance level by two orders of magnitude reduces the ensemble RMS by one order of magnitude. Kelly, Law, and Stuart [67] and Sanz-Alonso and Stuart [91] showed that, for linear models and in the absence of process noise and modeling errors, the long-term error and variance levels are directly linked to the measurement noise level.

(a) State estimate error   (b) Ensemble RMS.
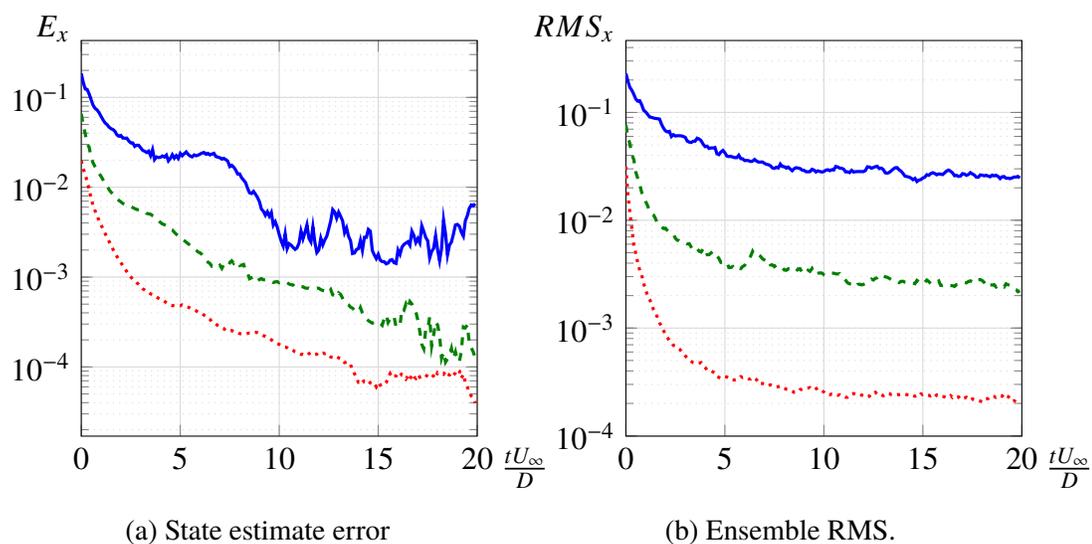
Figure 3.3: Estimator performance for distinct measurement noise levels applied to the problem of the flow past a circular cylinder at $Re = 100$ using velocity measurements and 16 ensemble members. (——), (– – –) and (⋯⋯) correspond, respectively, to measurement noise covariance matrices R set to $10^{-2}I_p$, $10^{-4}I_p$ and $10^{-6}I_p$.

3.1.4   Analysis of the Representers

Figure 3.4 shows the measurement influence fields (representers) for different measurement locations, as defined in Eq. 1.58b. Each of these vectors represents the expected value of the cross correlation between the state and each of the measurements. It comprises two pieces of information: its magnitude provides a comparative measure on how strongly the data gathered from each of the sensors contributes to

the correction, and its shape shows where in the flowfield these corrections take place. The highest values are achieved for measurements taken 2 to 3 diameters downstream of the cylinder. The obvious conclusion is that measurements taken at points where there is no variation amongst the ensemble members are useless. Interestingly, the measurement locations with higher representer magnitudes is the ones closer to the so-called *wavemaker region* [92], as shown in Fig. 3.4f. Note, however, that this analysis is made *a posteriori*. As such, it can be used to evaluate the relative relevance of the data obtained from each of the sensors, and propose changes to the disposition of sensor locations.

3.2 Flow Past a NACA 0009 Airfoil at High Angle of Attack

Throughout this section, the flow past a NACA 0009 airfoil at $Re = 200$ and $AoA = 30°$ (see section 2.2.2 for more details) was used as test problem. The grid resolution was set to 50 points by diameter ($Re_\Delta = 4$), and pressure measurements were taken at 10 equidistant points on the surface of the airfoil every 0.05 convective time units. The RTPS multiplicative covariance inflation scheme (see section 1.5.4) with $\theta = 0.95$ was used. The ensemble size was set to 24, and is initialized using the POD-based scheme shown in section 1.5.2.

Figure 3.5 shows the estimated lift coefficient evolution for this baseline setup. The behavior of the ensemble variance, here represented by the width of the shaded area, can be analyzed using Eq. 1.58a. Before the first analysis step, the variance of the ensemble-predicted measurements $H_k \hat{C}_k H_k^T$ is much larger than the measurement noise level R_k (see vertical bar near the y-axis). As a consequence all ensemble members are strongly corrected toward the true solution in the first time step. At subsequent times, corrections are increasingly damped as the estimate converges to the true flow state and the perceived uncertainty of the estimated measurements decreases.

3.2.1 Effects of the Covariance Inflation Scheme

We now analyze the effect of using different choices of covariance inflation (CI) scheme. These schemes delay the decrease in the ensemble covariance (see Figs. 3.6c and 3.6d) and, as a consequence, achieve a smaller error than the one obtained without inflation (see Figs. 3.6a and 3.6b). Because a constant factor is used to inflate the covariance in the AA scheme, the error subspace spanned by the ensemble is preserved. The corrections in this case are structurally similar to those that would

(a)

(b)

(c)

(d)

(e)

(f) Wavemaker region. Reproduced from [92].
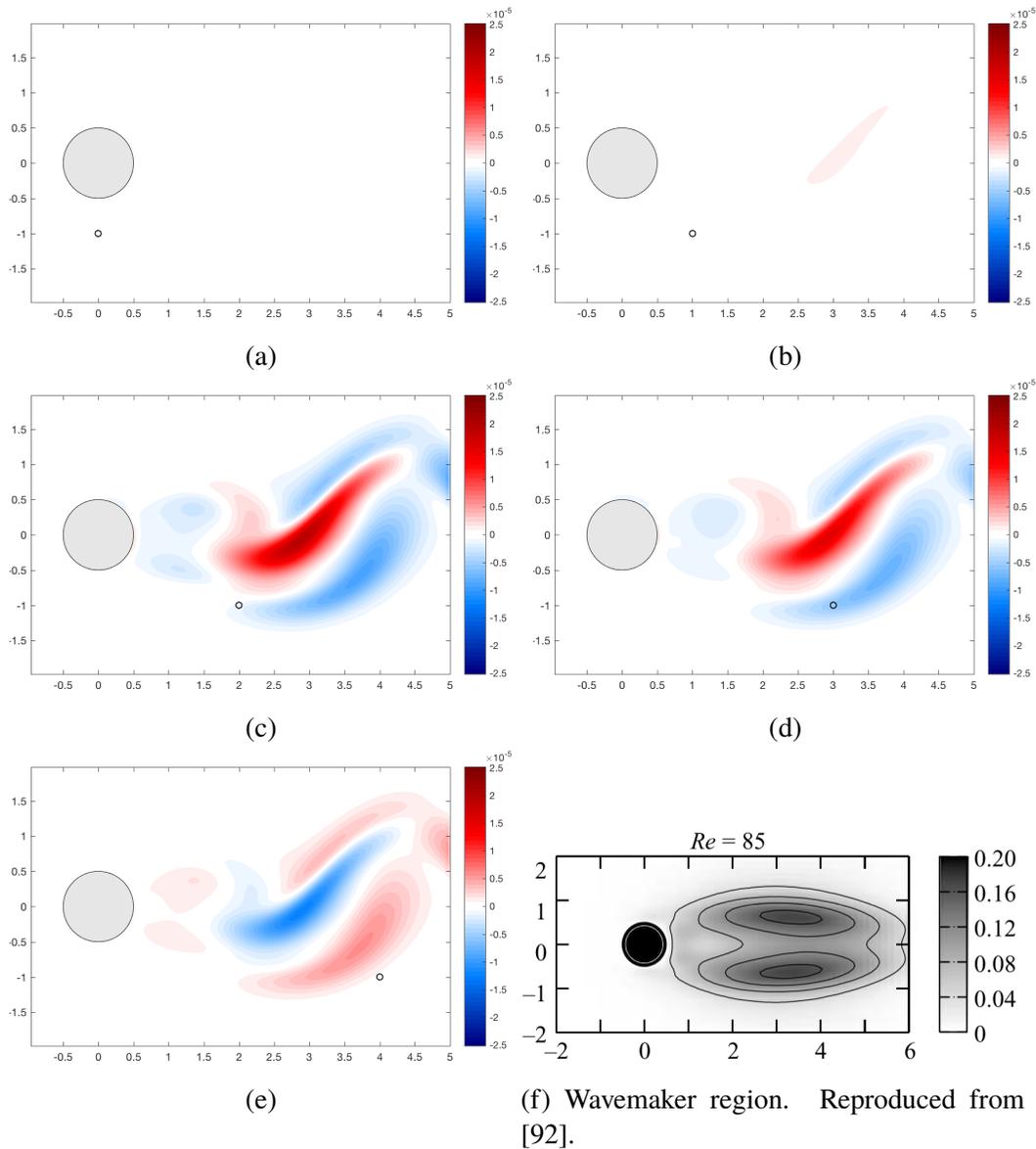
Figure 3.4: Measurement influence fields (representers) for the horizontal component of the fluid velocity at selected measurement locations. Ensemble size was 16 members and $R = 10^{-4}I_p$. All the figures have the same contour levels.

have taken place in the absence of inflation, but their amplitude is more aggressive. The RTPS scheme, on the other hand, is more complex. The inflation magnitude is local and dependent upon the previous analysis step (for $\theta = 1$, the forecast spread is recovered). Because of these characteristics, the RTPS scheme can in fact change the subspace spanned by the ensemble. This extra variability is possibly the reason for the distinct behavior at early times and may be even desirable for some applications.
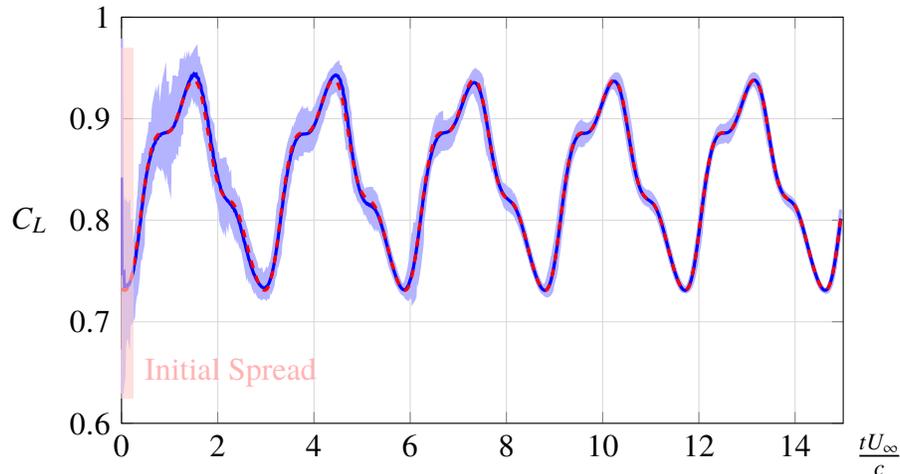
Figure 3.5: Estimated lift coefficient for the flow past a NACA 0009 airfoil using pressure measurements and the RTPS covariance inflation scheme ($\theta = 0.95$). The dashed line (- - -), the solid line (——), and the shaded area (▬) represent, respectively, the reference solution, the EnKF estimate, and the ensemble min-max envelope ($R = 10^{-4}I_p$). The bar on the vertical axis represents the initial ensemble spread.

3.2.2 Effect of the Number of Measurements: an Iterative Approach for Optimal Sensor Placement

The relative magnitude of the measurement influence fields (representers) are now used to iteratively search for the best locations for sensor placement. Starting with a simulation in which measurements are taken at all immersed boundary points, the following methodology was adopted:

1. A simulation is run for T convective time units. At each analysis step, the representers are sorted in descending order by their L_2 norms and the leading n are selected.

2. A histogram is used to represent how often each representer is selected during the simulation.

3. The n most frequent sensor locations according to the aforementioned histogram are selected.

4. A new simulation is run with only the selected sensor locations and the process is repeated until the desired number of sensors is achieved.
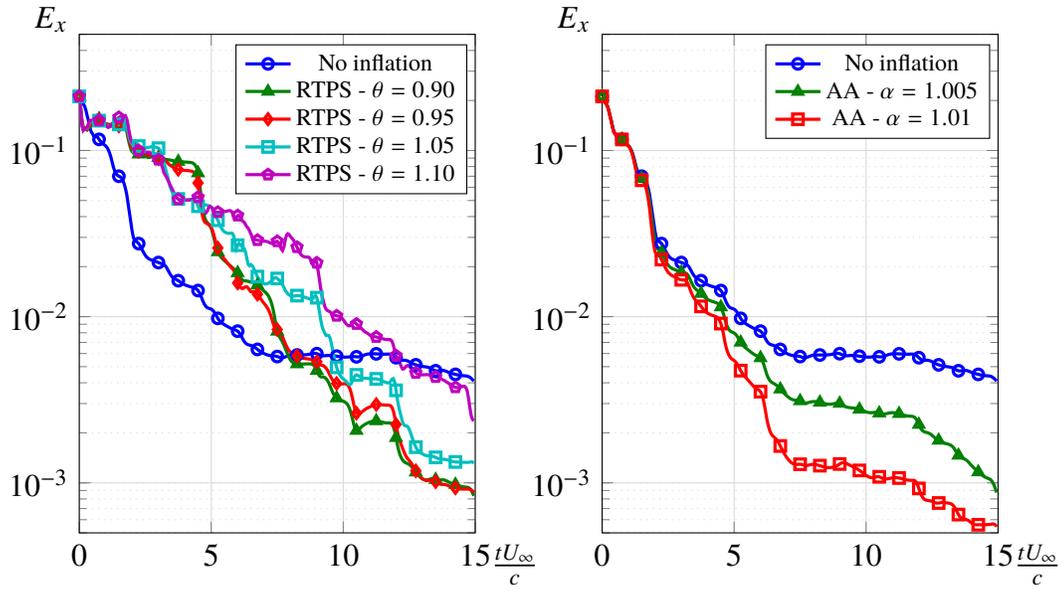
(a) State estimate error: RTPS scheme.

(b) State estimate error: AA scheme.
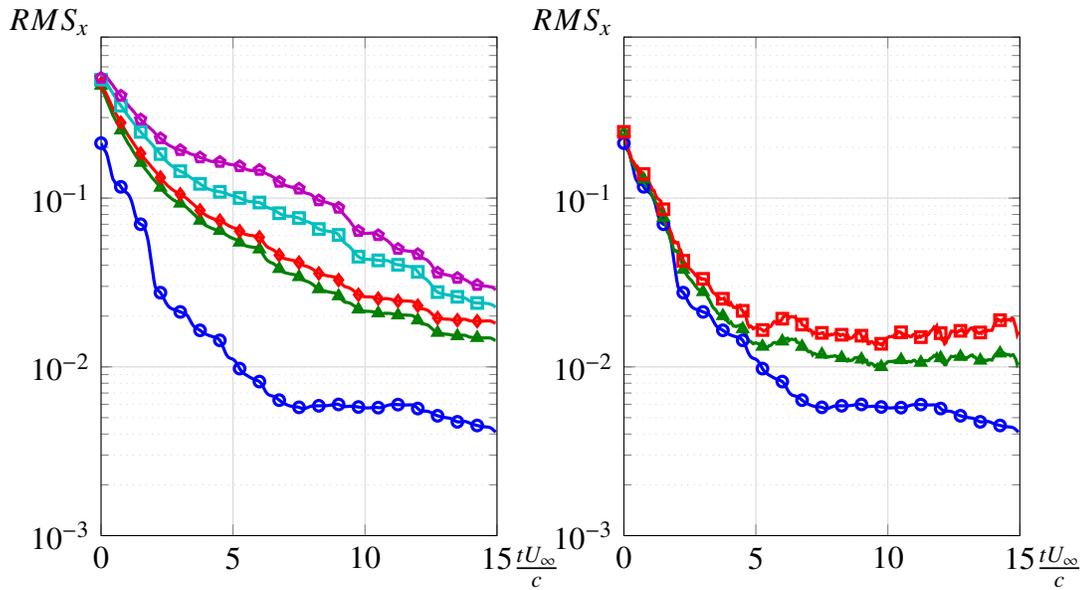
(c) Ensemble RMS: RTPS scheme.

(d) Ensemble RMS: AA scheme.

Figure 3.6: Estimator performance for different multiplicative covariance inflation schemes applied to the flow past a NACA 0009 at high angle of attack. Ensemble size is 24 and the sensor noise level is set to $R = 10^{-4}$.

The length of the simulation T must be chosen carefully. If too small a T is chosen, sensor locations more relevant to the filter transient behavior will be favored. On the other hand, a longer simulation length will favor the sensor locations associated with the filter long-term behavior and the initial filter performance may be hindered.

Here, T was set to 20 convective time units, which is about when the estimation statistics reaches a long-term stationary value.

Figure 3.7 shows a subset of the histograms obtained for the airfoil problem. Indexing starts at the trailing edge and increases as the airfoil surface is transversed in the counterclockwise direction. Starting with the first iteration, the regions close to the leading and trailing edges are favored. These regions play an active role in the formation and release of the vortices that characterize this particular flow and, therefore, pressure measurements taken at these locations should be especially effective to determining the state of the system. In fact, the estimator is able to track the vortex shedding phase with as few as 1 measurement located at either the leading or trailing edge.

(a) n=100

(b) n=50

(c) n=10

(d) n=2
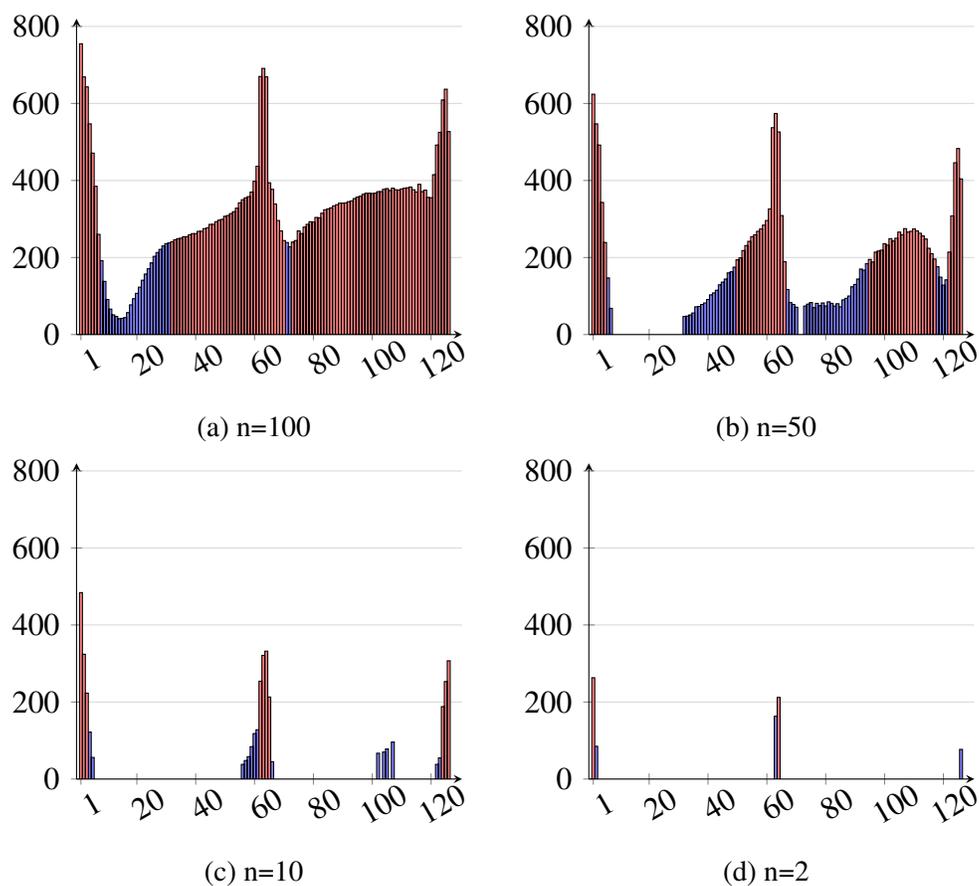
Figure 3.7: Histograms used in the iterative process of finding the optimal pressure sensor placement for the airfoil problem. Indexing starts at the trailing edge and increases as the airfoil surface is traveled in the counterclockwise direction. The blue bars (▌▌) represent the active sensors for that particular simulation, and the red bars (▌▌) represent the sensor locations selected for the next iteration.

The estimated state error, whose norm was previously presented in Fig. 3.6, is visualized at final instant of the estimation in Fig.3.8. Since no modeling errors are present, a global match is achieved even with local measurements after the estimator synchronizes with the reference solution.

(a) Estimated solution.

(b) Reference solution.
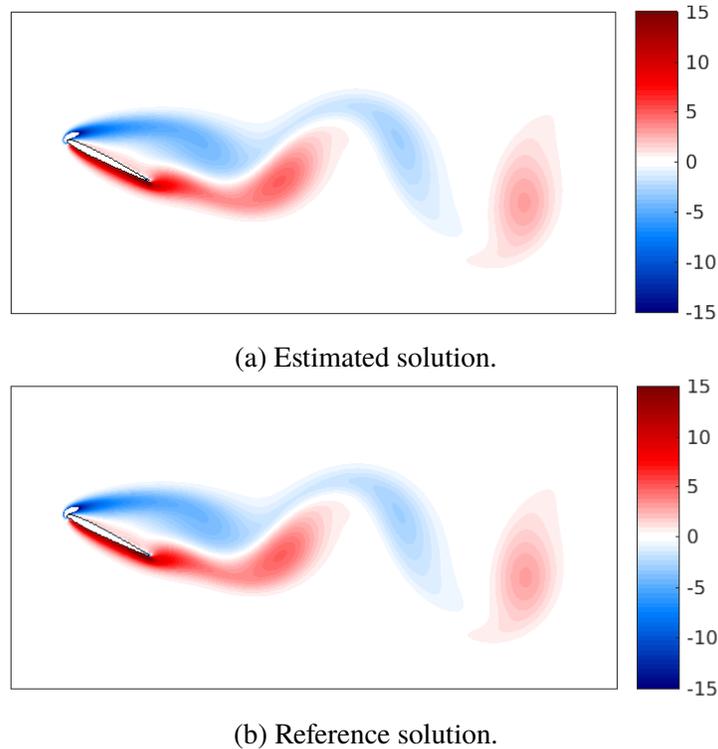
Figure 3.8: Estimated and reference vorticity field 20 convective time units after estimator initialization using a single pressure sensor optimally placed near the leading edge.

3.3 Flow Past an Inclined Flat Plate

In this last section, the flow past an inclined flat plate at $Re = 200$ and $AoA = 30°$ (see section 2.2.2 for more details) was used as test problem. The grid resolution was set to 50 points by diameter ($Re_\Delta = 4$), and pressure measurements were taken at 10 equidistant points on the plate (including leading and trailing edges) every 0.05 convective time units. The RTPS multiplicative covariance inflation scheme (see section 1.5.4) with $\theta = 0.90$ was used.

3.3.1   Computational Time Expenditure

As pointed out in the introduction, the EnKF methodology is highly parallelizable. The only serial step of the algorithm is represented by the evaluation of the ensemble statistics (or the perturbation matrices, when the ensemble transformation is employed), which requires gathering information from only ensemble members. Because we can choose between performing the matrix inversion in the Kalman gain calculation either in the measurement space or the ensemble space, the analysis step is usually relatively inexpensive[2].

Figure 3.9 shows how the computational time is distributed among the different filtering steps for a simulation ($q = 60$) using pressure measurements and the implicit linearization scheme. We separate the evaluation of measurements estimates from the two classic Kalman steps since this task is readily parallelizable at the ensemble level, but some researchers consider to be part of the analysis step. Note that the analysis step, the only part of the algorithm that is intrinsically serial, corresponds to a minor fraction of the time (about 2%). Because only 12 threads were used in this case, there is still space for a speed up in the ensemble level, not to mention that the forecast and observation models may be parallel in its own right.

3.3.2   Effect of Nonlinearity in the Measurement Function

We first analyze how the choice of the method to deal with the nonlinear observation function impacts the performance of the estimator in a perfect-model framework. Here, the reference solution was taken to be a simulation carried out in the same resolution level used by the estimator. Whenever the iterative method (Gauss-Newton) is used, the initial guess is obtained by the implicit linearization scheme. Since the maximum-likelihood estimate produced by the iterative method will only correspond to the minimum variance estimate in the case of a linear observation function, it is expected the posterior variance obtained by the iterative method will be larger than optimal. It was also noted that, at the initial assimilation steps, the iterative scheme achieves a reduction of the cost function of at least two orders of magnitude. As the estimation progresses and the state variance decreases, the magnitude of the corrections decreases and the cost improvement also becomes marginal (see Fig. 3.10a).

---

[2]Either because the computational cost associated with the forecast of a large ensemble is prohibitive or, as as happens in the present study, the size of the measurement vector is reduced.
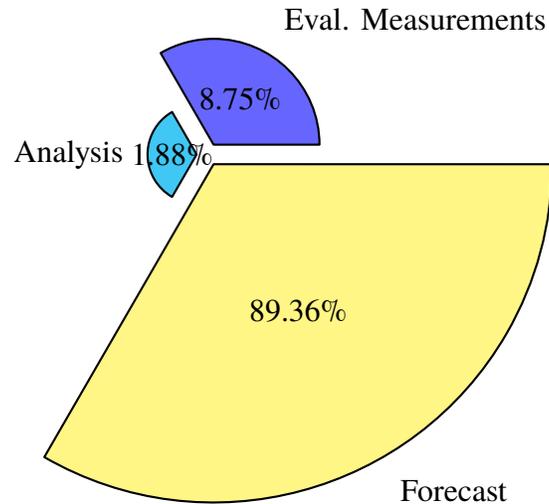
Figure 3.9: Time spent in each of the filtering steps for the EnKF simulation with pressure measurements, implicit linearization scheme, and $q = 60$. Twelve threads were used to split the computational load in the forecast step and when gathering measurements.

Despite its simplicity and lower cost, the implicit linearization scheme performs surprisingly well compared to the iterative method. Figure 3.10a shows that this scheme performs better, or at least similarly, than the iterative scheme for all the studied measurement noise levels (R). It is not clear how well the linearization about the ensemble mean is supposed to approximate the conditional mean (minimum error variance) of the posterior distribution, but results seem to indicate that it yields a better approximation than the iterative method.

In the absence of nonlinearities and modeling errors, the terminal state error level is supposed to decrease as the measurement noise level decreases (see [93]). Since errors due to the presence of nonlinearities in the observation function are expected to scale with $h_{xx}\hat{C}_k^{xx}$ [36], and Fig. 3.10b shows that lower R consistently produces smaller variances, the saturation of the state error at 10^{-3} for both the schemes is unexpected. Further study is necessary to investigate what source of error, independent of R or $\hat{C}_k$, is dominating the long-term behavior of the filter.

3.3.3 Optimality of the Analysis Step

If a Kalman filter works optimally, the time sequence of the innovation vectors $r_k = y_k - h(x_k)$ is white. The whiteness of this sequence is an indicative that the filter is able to extract all available information from the measured data and

(a) State error evolution.　　　(b) Observation error evolution.
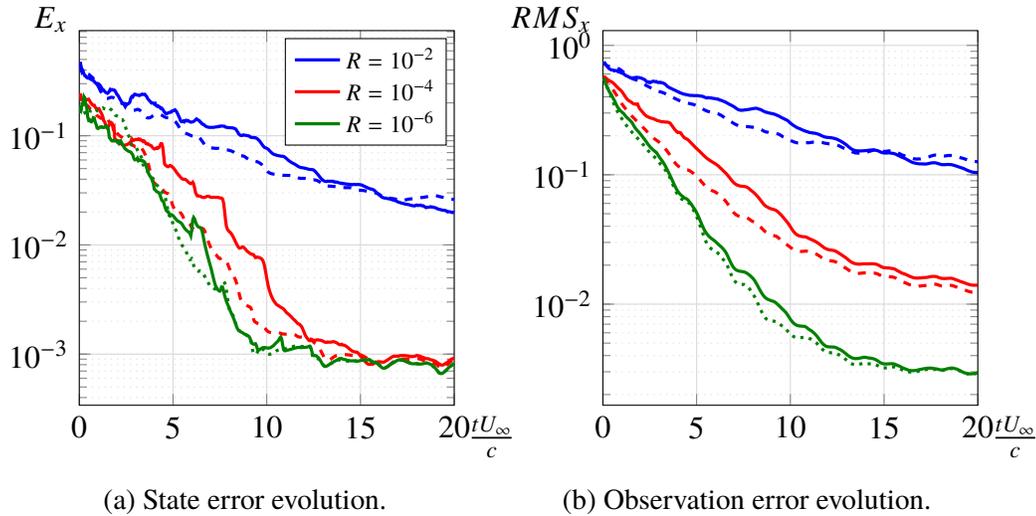
Figure 3.10: Effect of nonlinear observation model in the estimator performance in the perfect-model framework. Solid lines correspond to the iterative (Gauss-Newton) method and dashed lines, to the implicit linearization scheme.

incorporate it into the estimate. Therefore, the extent to which the innovation is white can be used as a measure of the filter optimality. The optimality of the EnKF will now be evaluated using both pressure and velocity measurements for the inclined flat plate case.

In order to verify the zero-mean requirement, we evaluate the expectation normalized innovation, which is defined as

$$\bar{r}^* = \mathbb{E}\left[\frac{r_k}{y_k}\right] , \qquad (3.1)$$

where the expectation is taken in time. Table 3.1 shows the values of $\bar{r}^*$ corresponding to the use of either pressure or velocity measurements for estimation window of 10 convective time units. If the whiteness hypothesis holds, the entry-wise sum of $\bar{r}^*$ is supposed to distribute as a Gaussian with zero mean and variance $tr(R)/N$, where $N = 10/0.05 = 200$. If r is white, there is a 13% chance for pressure measurements and a 73% chance for velocity measurements that we would measure $\sum \bar{r}_i^*$ this far from zero.

Table 3.1: Norm of the mean normalized innovation for different choices of measurements.

	Pressure	Velocity
$\sum \bar{r}_i^*$	0.3397%	0.1203%

In order to be a white sequence, in addition to having a zero mean, the sequence $r'_k = r_k - \mathbb{E}[r_k]$ should be uncorrelated in time. Figure 3.11 shows two examples of the autocorrelation function of such sequences for two measurement locations. Any values outside the dashed lines representing the confidence bounds should be considered significant because they are unlikely to occur just by chance. Thus, these results indicate that the innovation sequence is white for the velocity measurement but not for the pressure measurement, which exhibit some correlation for 1 lag.

In order to provide a more systematic way of accessing the whiteness of innovation vector, two tests proposed by Matisko and Havlena [94] will be used to demonstrate this hypothesis. They are based on the analysis of the autocorrelation function of the innovation sequence. For an innovation sequence of length N, we first compute the auto-covariance matrix [3]

$$M^*(\tau) = \mathbb{E}[r'_k r'^T_{k+\tau}] = \frac{1}{\sqrt{N(N-\tau)}} \sum_{k=1}^{N-\tau} r'_k r'^T_{k+\tau} , \qquad (3.2)$$

where $\tau \in \mathbb{N}$. The correlation function can be then obtained by appropriately normalizing the previous equation

$$M_{ij}(\tau) = \frac{M^*_{ij}(\tau)}{\sqrt{M^*_{ii}(0)M^*_{jj}(0)}} , \qquad (3.3)$$

where the subscripts ij denotes a specific entry in the respective matrix.

Test 1: Innovation Autocorrelation Function

If $r'_k \sim N(0, R)$, then we expect $M_{ii}(\tau) \sim N(0, 1/N)$ for $\tau > 0$, i.e., for each $\tau > 0$ and $1 \geq i \geq p$, the value of $\sqrt{N}M_{ii}(\tau)$ can be regarded as being independently sampled from a Gaussian distribution with zero mean and unitary variance. Therefore, the auxiliary function

$$\Phi_j(\tau_{max}) = N \sum_{\tau=1}^{\tau_{max}} \left(M_{jj}(\tau)\right)^2 , \qquad (3.4)$$

where τ_{max} is the maximum lag being taken into consideration, is expected to distribute as a chi-squared distribution with τ_{max} degrees of freedom. The innovation sequence r'_k is considered white if, for $1 \leq \tau \leq \tau_{max}$,

$$\Phi_j(\tau) < q^\tau_{5\%} , \qquad (3.5)$$

---

[3]Note that we have used a non-conventional pre-multiplying constant for the auto-covariance matrix definition. Its value is chosen so that $M(\tau)$ are i.i.d for $\tau > 0$.
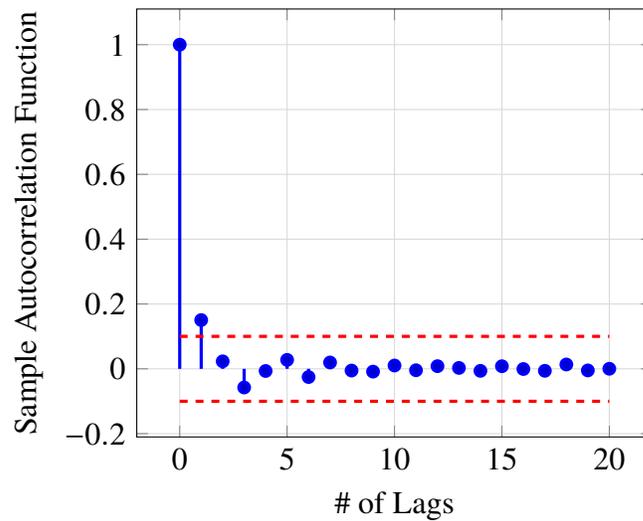
(a) Autocorrelation function for the innovation corresponding to the pressure measured at the trailing edge.
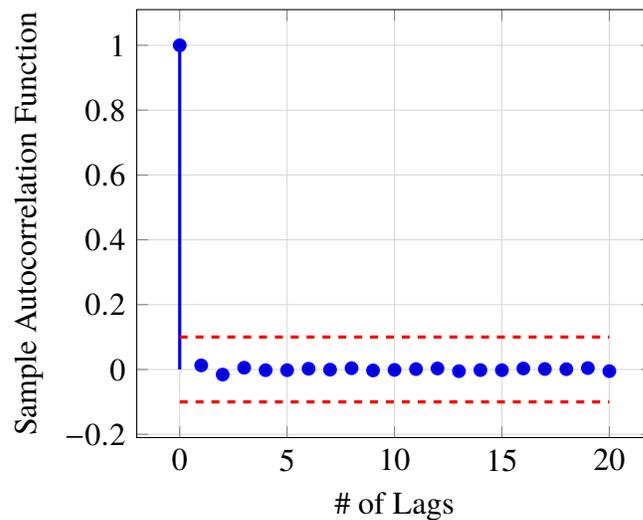
(b) Autocorrelation function for the innovation corresponding to horizontal velocity measured one diameter downstream of the flat plate.

Figure 3.11: Examples of the sample autocorrelation function for different measurement locations. Horizontal dashed lines correspond to the associated upper and lower confidence bounds, respectively.

where $q_{5\%}^{\tau}$ is the 5th percentile of the chi-squared distribution with τ degrees of freedom. This criteria can be interpreted as follows: left to random chance, there is only 5% chance of observing values bellow $q_{5\%}^{\tau}$, and, therefore, actually observing this occurrence is indicative that the corresponding sequence is white.

Figure 3.13 shows the curves for $\Phi(\tau)$ corresponding to the horizontal and vertical components of the velocity measurements. The fact that the curves don't grow indefinitely but saturate after a finite number of lags is an indication that the sequences are not periodic. Moreover, most of the curves remain below the dashed line corresponding to the 5%-significance interval, indicating that the data corroborates the whiteness hypothesis for the innovation sequences corresponding to those measurement locations. However, there is strong evidence of correlation for the horizontal velocity measured 4 chords downstream of the plate, and a weaker evidence for the vertical velocity two chords downstream in the top row.

Comparing Fig. 3.12 with Fig. 3.13, one can clearly see that the velocity measurements are more efficiently assimilated to the estimate than the pressure measurements. This loss of optimality in the pressure measurement case is not surprising given the fact that the pressure observation function is nonlinear. It is also noteworthy the fact that the pressure locations in the second half (close to the trailing edge) of the plate display stronger evidence of not being white.

Test 2: Best Auto-regressive Model

If r'_k is a white sequence, then the best auto-regressive model to represent the data should be of zeroth order. The general expression for an ARn model is

$$r'_k = \sum_{i=1}^{n} \Lambda_i r'_{k-i} + \epsilon_k \ , \tag{3.6}$$

where $\Lambda_i \in \mathbb{R}^{p \times p}$ are the model coefficients, and ϵ_k is a noise process. The auto-covariance matrices computed for the previous test can be used to obtain the appropriate coefficients Λ_i by solving the multivariate Yule-Walker equations:

$$\begin{bmatrix} M^*(0) & M^*(1) & \cdots & M^*(n-1) \\ M^*(1) & M^*(0) & & M^*(n-2) \\ \vdots & & \ddots & \vdots \\ M^*(n-1) & M^*(n-2) & \cdots & M^*(0) \end{bmatrix} \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \\ \vdots \\ \Lambda_n \end{bmatrix} = \begin{bmatrix} M^*(1) \\ M^*(2) \\ \vdots \\ M^*(n) \end{bmatrix}$$

$$\Longleftrightarrow \mathbb{M}\Lambda = b \ . \tag{3.7}$$
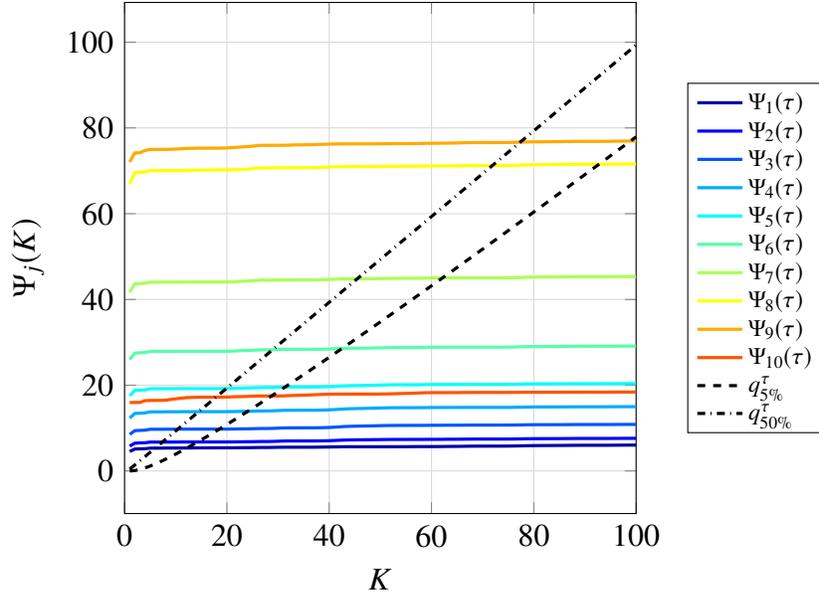
Figure 3.12: Optimality of the EnKF with pressure measurements as perceived from the decay of the innovation autocorrelation function. Measurement points are numbered from the leading edge to the trailing edge.

The residual covariance matrix, whose norm is a measure of goodness of fit of the corresponding ARn model, is given by

$$S^2 = M^*(0) - \sum_{i=1}^{n} \Lambda_i M^*(i) = M^*(0) - \Lambda^T \mathbb{M} \Lambda \ . \tag{3.8}$$

Since norm of the residual is expected to be non-increasing as one increases the order of the auto-regressive model (as a lower-order model is a special case of higher-order models), it is necessary to use some penalizing criteria to avoid over-fitting. Following Pukkila and Krishnaiah [95], we adopt the BIC criterion that uses a penalty term of the form $g(n) = np^2 \log N$. Therefore, finding the optimal order corresponds to solving

$$\begin{aligned} n_{opt} &= \underset{0 \leq n}{\operatorname{argmin}} \Phi(n) \\ &= \underset{0 \leq n}{\operatorname{argmin}} \ \left[N \log \left(\det(S^2) \right) + np^2 \log N \right] \ . \end{aligned} \tag{3.9}$$

Figures 3.14 and 3.16 display the values of the cost function $\Phi(n)$ for each of the measurements locations. The auto-regressive orders that correspond to the minimum values of the cost function for each of the measurement locations are presented in Fig. 3.15 and 3.16 for pressure and velocity measurements, respectively. For the
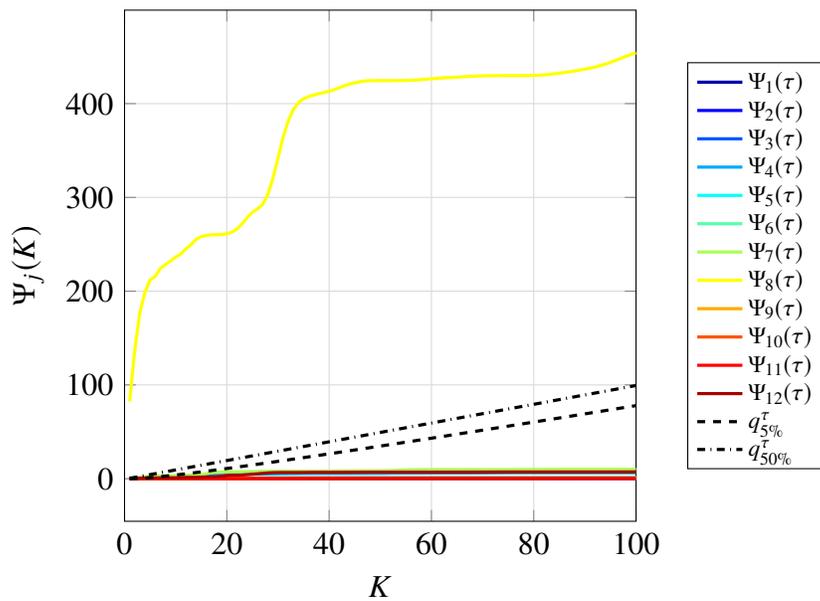
(a) Horizontal velocity measurements.
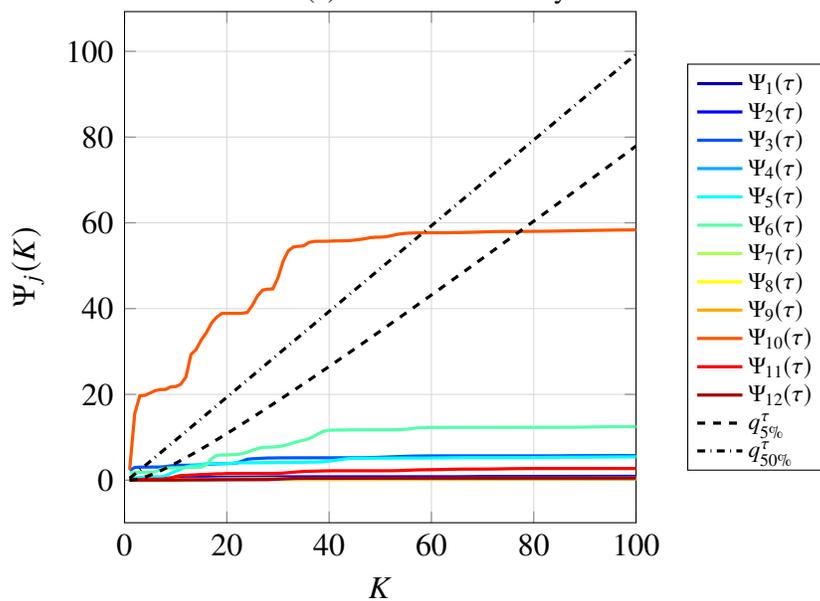
(b) Vertical velocity measurements.

Figure 3.13: Optimality of the EnKF with velocity measurements as perceived from the decay of the innovation autocorrelation function. Measurement points are numbered from left to right, and from the bottom to the top.
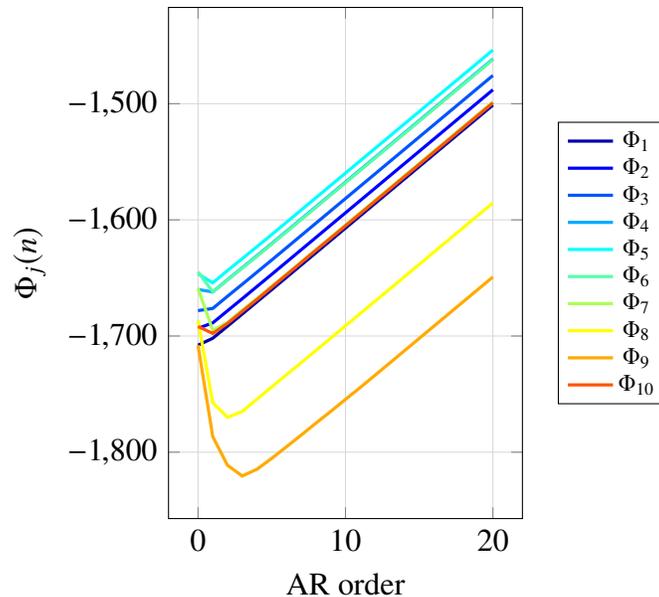
Figure 3.14: Optimality of the EnKF with pressure measurements as perceived by finding the best AR model that represents the innovation sequence.

most part, the results from the second test corroborate the ones obtained from the first test. Consistently with the previous observations, the innovation sequences associated with pressure measurement taken in the second half of the plate exhibit wider correlation, with 3 being the maximum number of lags. As for the velocity measurement points, most innovation sequences are deemed white, with the single exception being the horizontal measurement 4 chords downstream of the plate, where a one-lag correlation was indicated. The only conflicting result is relative to the velocity location in the upper row that, having been flagged as non-white by the first test, is now cleared by the second test. In general lines, both tests agree with the assessment that the filter is able to assimilate velocity measurements more efficiently.

3.3.4 Comparison to the 3D-EnVar

The 3D-EnVar is a variant of the 3D-Var (presented in section 1.4.2) in which we use an ensemble to build a low-order representation of the constant matrix Σ.

$$X = \begin{bmatrix} x^{(1)} & x^{(2)} & \cdots & x^{(N)} \end{bmatrix}$$

$$\Sigma = \frac{\alpha}{N-1} \left(X - \frac{1}{N} X \mathbb{1} \mathbb{1}^T \right) \left(X - \frac{1}{N} X \mathbb{1} \mathbb{1}^T \right)^T = \alpha A A^T , \qquad (3.10)$$
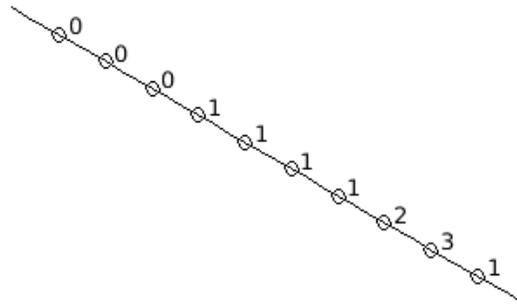
Figure 3.15: AR order attributed by the second test to each of the pressure measurement points.
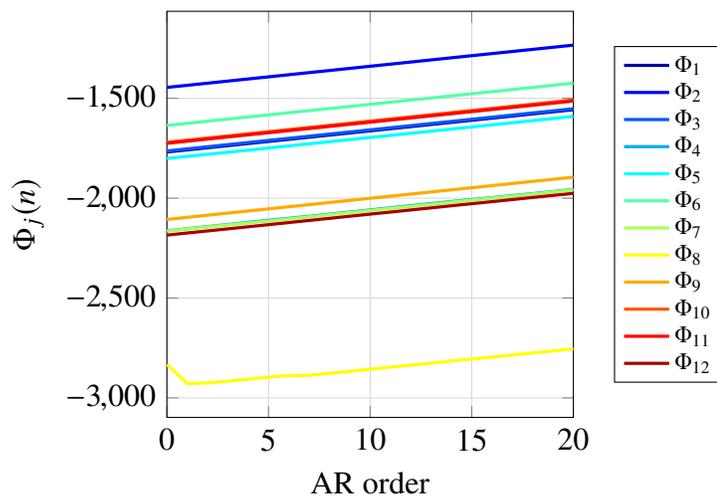
(a) Horizontal velocity measurements.
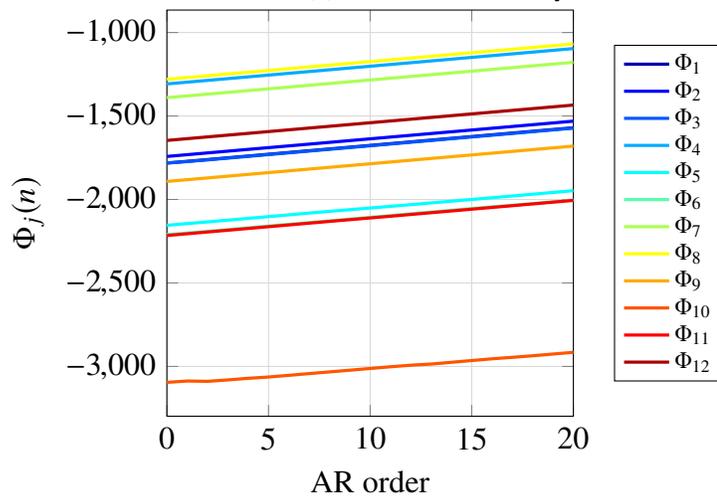
(b) Vertical velocity measurements.

Figure 3.16: Optimality of the EnKF with velocity measurements as perceived by finding the best AR model that represents the innovation sequence.
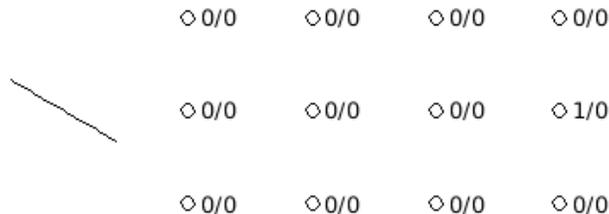
Figure 3.17: AR order attributed by the second test to each of the velocity measurement points. First and second values correspond, respectively, to the horizontal and vertical component of the velocity vector.

where the data matrix X is built of a sequence of snapshots sampled during a long integration of the model. Since an ensemble of model states will naturally have a variance that is too large to represent the actual error in the model forecast, a parameter α is used to scale it. In other words, we can understand α as a covariance deflation parameter. This methodology provides a consistent way of building the weighting matrix for the 3D-Var, and also allow us to understand the impact of the covariance inflation alone. As a consequence of this particular choice for Σ, the search for a minimizer of the associated cost function is also restricted to the low-dimensional affine subset generated by the prior estimate $\hat{x}$ and the perturbation matrix A.

Figure 3.18 shows the performance of the 3D-EnVar scheme for different values of the parameter α. Measurements consist of the velocity components sampled at 15 different locations (c.f. Fig. 2.7) and have noise covariance matrix $R = 10^{-4}I_p$. Note that the presence of the scaling factor is able to improve the performance by almost two orders of magnitude (for $\alpha = 0.25$). For higher values of α, the action of the model prediction as a regularizing term is minimized, and the long-term error is the poorest. In fact, in the limit we would be substituting the prior estimate $\hat{x}$ by the solution of the inverse problem $Hx = y$. This is a ill-posed problem (since $p << n$) that can be very sensitive to the measurement errors. On the other hand, when α is too small, the relative reliability of the prior is larger than any information that can be extracted from the measurements, and, consequently, any corrections will be small (which explains the slower convergence rate experienced by the $\alpha = 0.001$ case). The optimal case, which merges the right amount of information from the model forecast and the measurements, corresponds to intermediate values. Extrapolating

this idea, one can understand why the underestimation of the system variance, which commonly takes place when small ensemble are used in the EnKF, is detrimental.

(a) State estimate error.

(b) Observation estimate error.
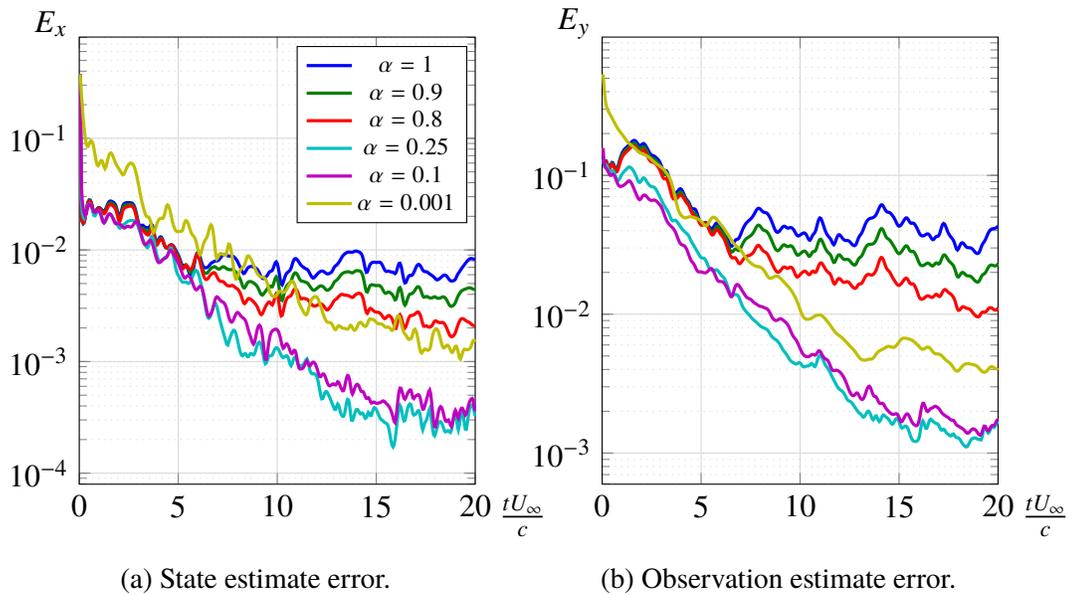
Figure 3.18: 3D-EnVar estimator performance for different values of α applied to the problem of the flow past a flat plate at $Re = 200$ and $AoA = 30°$ using velocity measurements. Measurement error level is set to $R = 10^{-4}I_p$.

Even though the 3D-EnVar is much less expensive than the EnKF, it has two main disadvantages. First, its performance is highly dependent on the proper choice of Σ and α. On the other hand, the EnKF performance is fairly robust with respect to choices of initial ensemble (c.f. section 1.5.2) and covariance inflation magnitude (c.f. section 3.2.1). Second, it does not provide any information on the estimation error. Although having its magnitude underestimated by the EnKF, a fair amount of information can be obtained from the structure of the estimated covariance matrix.

3.4 Summary

In this chapter, we sought to understand how different elements of the algorithm, such as measurement noise levels and covariance inflation schemes, impact the performance of the estimator. We analyzed approximate schemes that generalize the EnKF framework to deal with nonlinear observation functions, and showed evidence that they perform less optimally then when linear functions are used. We showed a correspondence between the representers field and regions of high structural sensitivity of the flow, and used that information as guidance for optimal sensor placement.

Although understanding the characteristics of an EnKF estimator in a perfect-model framework is important to obtain a better understanding of the filter dynamics, practical models are rarely able to exactly capture all the features of a dynamical system. In the next chapter, we take our first step into dealing with model inaccuracies by analyzing how the EnKF framework can be used to estimate model parameters.

Chapter 4

ESTIMATION WITH UNCERTAIN PARAMETERS

In practical applications, the success of the estimator is contingent on the accuracy of the model chosen to represent the dynamics. Imprecise or misspecified boundary conditions or incorrect a priori assumptions about flow parameters such as the Reynolds or Mach number can result in modeling errors that, left unaccounted for, can severely impair the estimator capabilities of accurately tracking the trajectory of the system. Examples of such parameters includes boundary conditions or flow parameters such as the Reynolds number.

If any further information about the sources of these errors is either unknown or ignored, a first strategy is what we call the *agnostic approach*. Following this approach, no modification to the model itself is proposed and multiplicative covariance inflation alone is used to reduce the perceived reliability of the estimator model in comparison with that of the data gathered from sensors[1]. Recall that inflating the ensemble covariance is equivalent to weighting the actual measurements to the detriment of the ones predicted by the estimator internal state [67], allowing more aggressive corrections to the state and enhancing the estimator responsiveness. This approach, however, has limitations. Artificially driving the state toward the measurements usually leads to larger state errors. If the modeling errors are large enough, a closer match in terms of observations may lead to completely unphysical states.

Provided that modeling errors can be tracked to theirs sources and suitable parametric models can be designed to represent them, an alternative strategy is possible. Following the *augmentation approach*[96], the new parameters are added to the state vector and the new augmented state is estimated using the EnKF framework.

In this chapter, we investigate the applicability of these two strategies to the problem of estimating freestream velocity pertubations and uncertain Reynolds number from pressure measurements taken on the surface of the immersed body.

---

[1]The multiplicative covariance inflation here can be understood as a process noise whose covariance matrix is a multiple of the state covariance matrix.

4.1 Freestream Velocity Perturbation

In order to investigate the capabilities of the agnostic approach, the basic framework introduced in section 3.2 is maintained but perturbations are introduced to the freestream velocity of the nature run, affecting both the forecast and measurement evaluations. The reference solution is randomly perturbed by setting the freestream velocity to

$$U_\infty(t) = 1 + 0.1\mathcal{F}_1(\xi(t)) , \qquad (4.1)$$

where $\xi(t)$ is a random Gaussian noise sequence and $\mathcal{F}_1(\cdot)$ is a 8th-order Butterworth filter with a predetermined cutoff reduced frequency (fc/U_{ref}). Initially, we adopt the agnostic approach and no modification to the estimator model is made other than increasing the multiplicative covariance inflation parameter. The RTPS inflation scheme is chosen over the AA scheme in order to introduce some variability to the subspace (spanned by the prior ensemble) from which corrections are drawn at each analysis step.

Figure 4.1 shows the ensemble lift coefficient evolution for two different choices of covariance inflation magnitude and ensemble size. Increasing the former leads to a more aggressive correction behavior that tends to yield noisier estimates. To some extent, this tendency can be mitigated with a larger ensemble size. A combination of both strategies usually leads to enhanced tracking capabilities.

Alternatively, having the augmentation approach in mind, we propose to model the freestream velocity as exhibiting a linear behavior between two consecutive analysis steps, such that the state vector is then augmented [96] with the inclusion of the freestream perturbation ξ and its time derivative $\dot{\xi}$. The dynamic model for them is a simple integrator with the derivative being propagated as a constant, as shown in Eq. 4.3.

$$U_\infty(t_k) = 1 + \xi_k \qquad (4.2)$$

$$\hat{z}_k = \begin{bmatrix} \hat{x}_k \\ \hat{\xi}_k \\ \hat{\dot{\xi}}_k \end{bmatrix} = \begin{bmatrix} f(x_{k-1}) \\ \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_{k-1} \\ \dot{\xi}_{k-1} \end{bmatrix} \end{bmatrix} + \tilde{\mu}_k = \tilde{f}(z_{k-1}) + \tilde{\mu}_k , \qquad (4.3)$$

where Δt is the time interval between two data assimilations, $\tilde{\mu}_k \sim N(0, \tilde{Q})$ is a Gaussian-distributed augmented noise vector whose associated noise covariance

(a) Ensemble size 24 and $\theta = 1.05$. (b) Ensemble size 48 and $\theta = 1.20$.
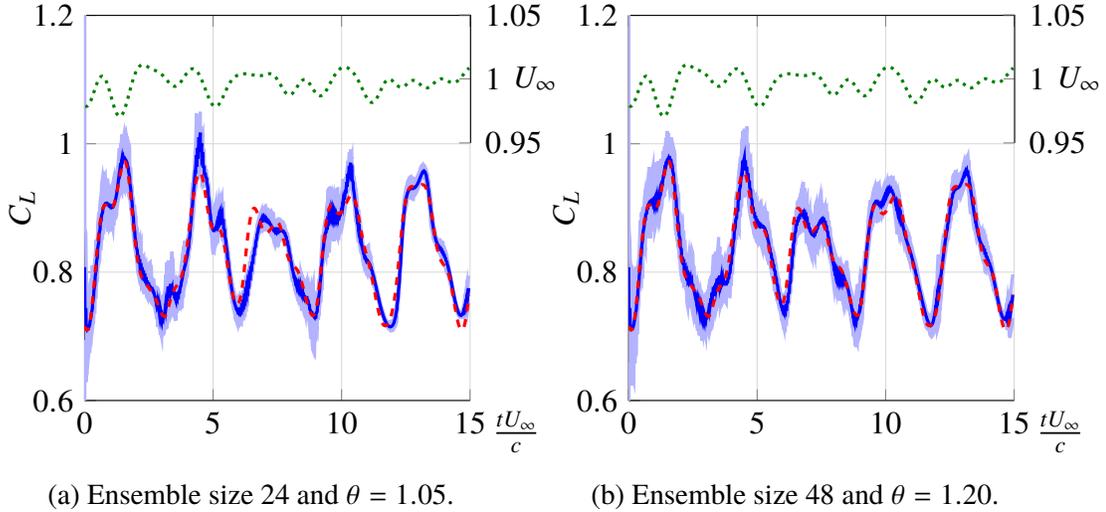
Figure 4.1: Estimated lift coefficient for airfoil with randomized freestream velocity using the RTPS covariance inflation scheme ($R = 10^{-4} I_p$). The dashed line (- - -), the solid line (——), and the shaded area (▪) represent, respectively, the reference solution, the EnKF estimate, and the ensemble min-max envelope. The dotted line (·····) represents the actual freestream velocity that the airfoil is subjected to.

matrix is given by

$$\tilde{Q} = \begin{bmatrix} Q & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma \end{bmatrix}. \tag{4.4}$$

This approach of simultaneously estimate both state and parameters is often referred to as *joint estimation*, and has the advantage of taking into account the cross-correlation between the state and the parameters.

First we analyze the performance of the estimator when the airfoil perform a smooth deceleration to 75% of its initial value over 13.3 convective time units so that 87.5% is achieved after 20 convective time units of the beginning of the simulation. The deceleration profile is given by

$$U_\infty(t) = 1 - \frac{\alpha}{2} \text{erfc}\left(-\frac{\sqrt{\pi}\beta}{\alpha}(t - t_0)\right), \tag{4.5}$$

where $\text{erfc}(t)$ is the complementary error function, α is the saturation value, β is the slope at $t = t_0$, the reference time at which the perturbation reaches half of its saturation value.

Figure 4.2 shows some interesting features of the EnKF. Because the IB forces acting on the body are a function of the vorticity field, and the freestream velocity

and acceleration, the lift (the sum of the vertical components of these forces) is a good global performance indicator for this estimation problem. For early times, the C_L ensemble variance is larger than the perfect model case due to extra degree of uncertainty added by the presence of a perturbation of unknown magnitude, but the estimator is able to obtain the right phase after just 2 convective time units and most of the perturbation tracking is performed while the estimator has already a good estimate for the state of the system.

Figure 4.3 shows how the parameters estimate evolves with a increasing inflation magnitude. For early times, the poorer estimates are consistent with the transient behavior that was observed in Fig. 4.2. The impact of the covariance inflation parameter is more pronounced in the acceleration, in which higher values lead to a better overall tracking but noisier initial estimates.
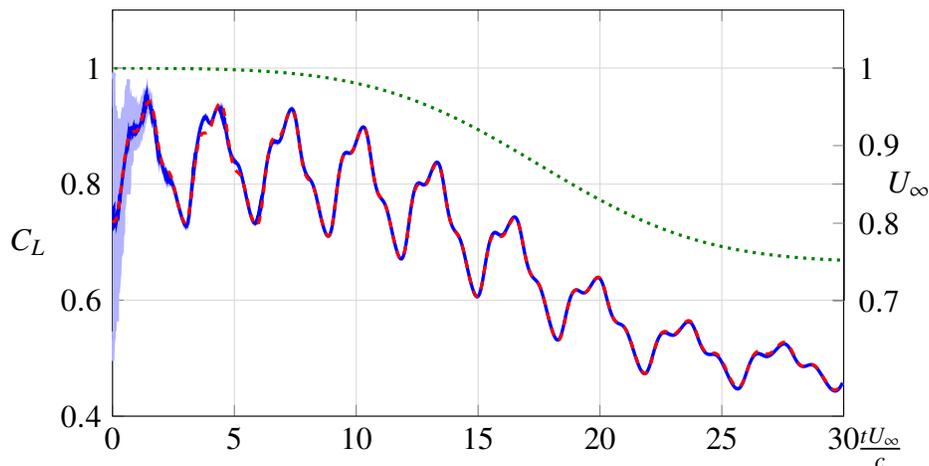
Figure 4.2: Estimated lift coefficient for decelerating airfoil (75% of its initial velocity over 13.3 convective time units) using the RTPS ($\theta = 0.9$) covariance inflation scheme and $R = 10^{-6}I_p$. The dashed line (- - -), the solid line (——), and the shaded area (▬) represent, respectively, the reference solution, the EnKF estimate, and the ensemble min-max envelope. The dotted line (·····) represents the actual freestream velocity that the airfoil is subjected to.

If the same deceleration is performed over a single convective time unit, keeping all others parameters the same, the performance of the estimator degrades (see Fig.4.4). The reason is that the freestream velocity only changes after 17 convective time units after the beginning of the simulation, when the estimator internal variance already reached its lower limit and the responsiveness of the estimator to external changes is already low (dashed line in Fig. 4.6). Therefore, it is necessary that the innovation norm grows before the estimator responds accordingly, leading to

(a) Free stream velocity perturbation.

(b) Free stream velocity perturbation derivative.
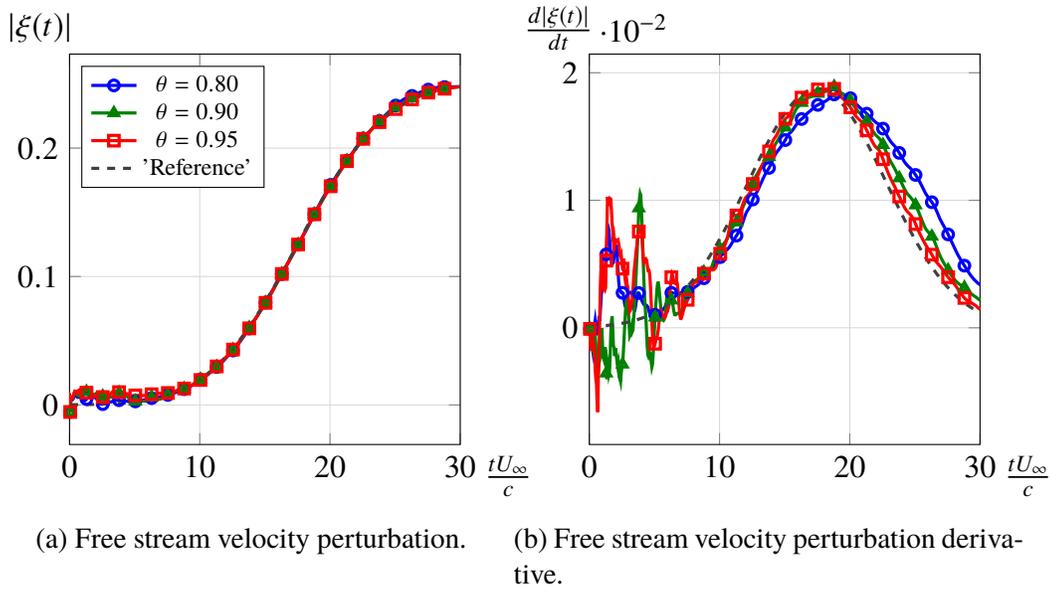
Figure 4.3: Joint parameter/state estimation for an airfoil decelerating to 75% of its initial value over 13.3 convective time units. The measurement noise level is set to $R = 10^{-6}I_p$.

the significant response delay observed in Fig. 4.4. Increasing the multiplicative covariance inflation would help to delay the variance decay, but its effectiveness will be directly linked to the size interval between the beginning of the estimation and the deceleration event. Also, a multiplicative covariance inflation that is too high will lead to noisier estimates for early estimation times.

Alternatively, one could stochastically force the parameter dynamics to keep their variance in a level that is high enough to ensure responsiveness. For this particular problem, we choose to force the disturbance derivative with a zero-mean Gaussian-distributed random quantity with a variance level one order of magnitude below the maximum expected derivative level so that the parameters ensemble covariance is kept above a minimum threshold (see solid line in Fig 4.6). Figure 4.5 shows that this addition enhances the estimator tracking performance.

Figure 4.8 shows how the magnitude of the variance of the process noise added to the acceleration parameter impact the error in the freestream perturbation estimation. The error in both plots is normalized by the maximum perturbation encountered within the time interval in consideration. Note that increasing the uncertainty perceived by the estimator leads to improved tracking performance. As another consequence of this added uncertainty, the min-max envelope for the C_L estimate also expands, as presented in Fig. 4.5.
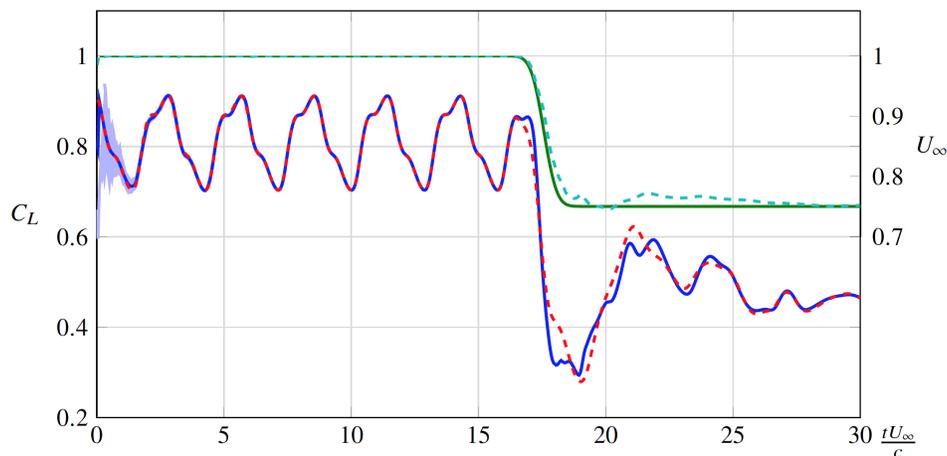
Figure 4.4: Estimated lift coefficient for decelerating airfoil (75% of its initial velocity over 1 convective time units) using the RTPS ($\theta = 0.9$) covariance inflation scheme and $R = 10^{-6}I_p$. The dashed line (- - -), the solid line (——), and the shaded area (▬) represent, respectively, the reference solution, the EnKF estimate, and the ensemble min-max envelope. The dotted line (⋯⋯) represents the actual freestream velocity that the airfoil is subjected to.
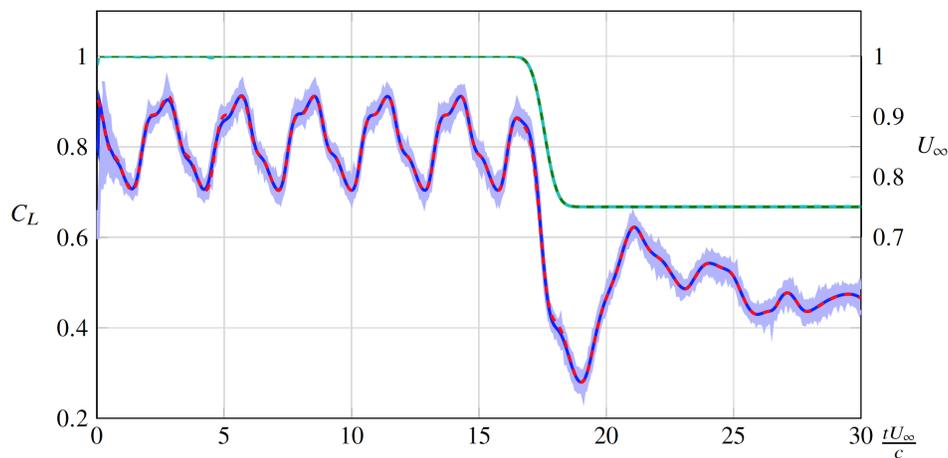
Figure 4.5: Estimated lift coefficient for decelerating airfoil (75% of its initial velocity over 1 convective time units) using the RTPS ($\theta = 0.9$) covariance inflation scheme and $R = 10^{-6}I_p$, and explicit stochastic forcing to the parameters dynamics ($\sigma = 10^{-8}$). The dashed line (- - -), the solid line (——), and the shaded area (▬) represent, respectively, the reference solution, the EnKF estimate, and the ensemble min-max envelope. The dotted line (⋯⋯) represents the actual freestream velocity that the airfoil is subjected to.

(a) Free stream velocity perturbation.

(b) Free stream velocity perturbation derivative.
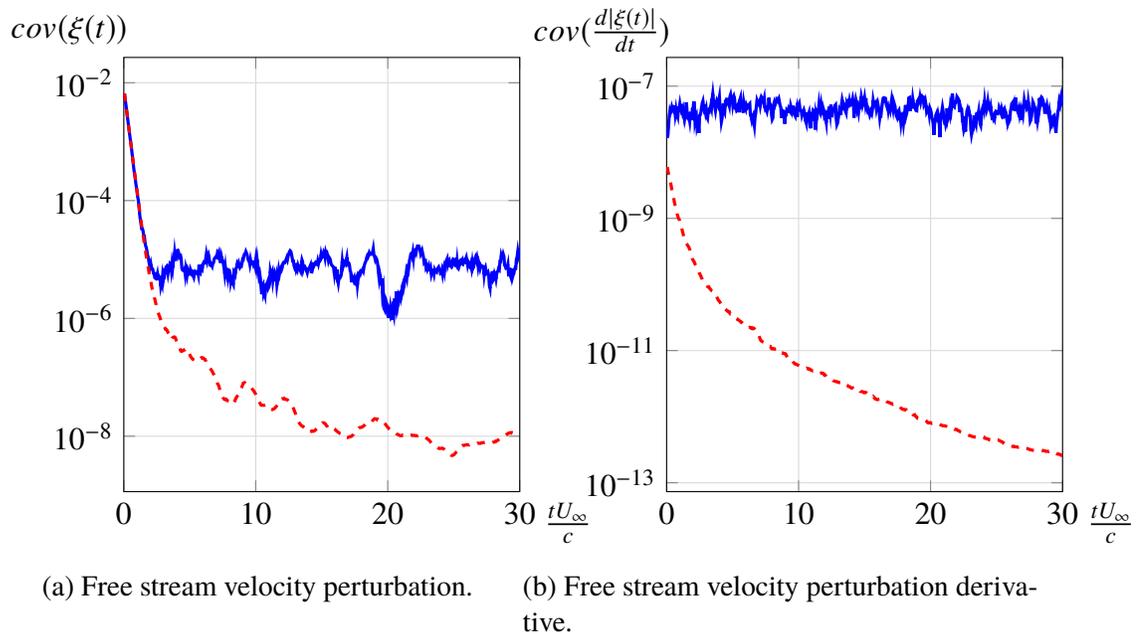
Figure 4.6: Evolution of the estimated parameters covariance for an airfoil decelerating to 75% of its initial value over one convective time unit. The measurement noise level is set to $R = 10^{-6}I_p$ and the RTPS inflation level is set to $\theta = 0.90$. The dashed line (- - -) corresponds to the unforced simulation and the solid line (—) represents the estimation dynamics with the explicit stochastic forcing in the perturbation derivative dynamics (noise covariance level 10^{-8}).

(a) Free stream velocity perturbation.

(b) Free stream velocity perturbation derivative.
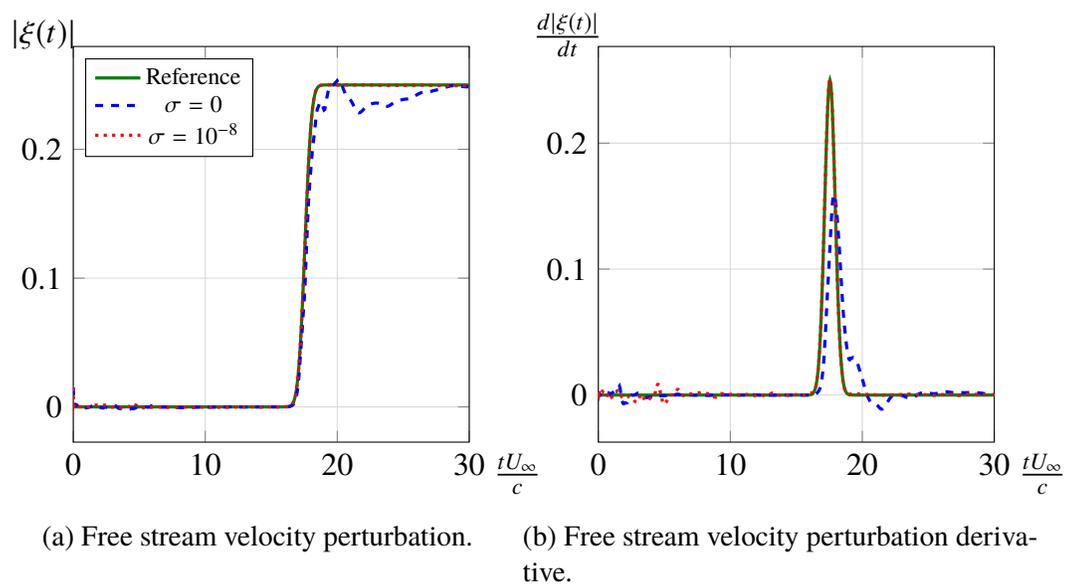
Figure 4.7: Joint parameter/state estimation for an airfoil decelerating to 75% of its initial value over one convective time unit. The measurement noise level is set to $R = 10^{-6}I_p$ and the RTPS inflation level is set to $\theta = 0.90$.
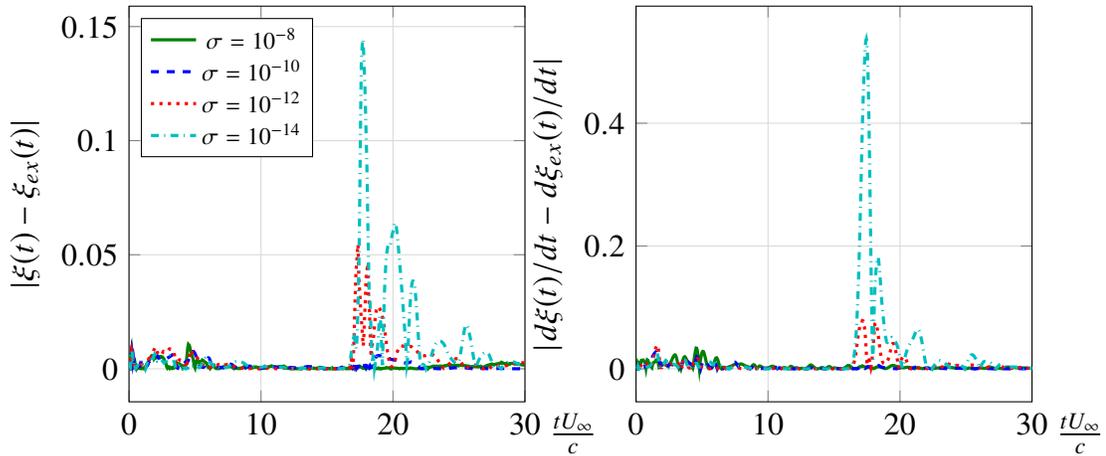
(a) Free stream velocity perturbation.

(b) Free stream velocity perturbation derivative.

Figure 4.8: Error in the estimation of the freestream parameters for an airfoil decelerating to 75% of its initial value over one convective time unit. The measurement noise level is set to $R = 10^{-6}I_p$ and the RTPS inflation level is set to $\theta = 0.90$. In both plots the error is normalized by the maximum perturbation in the interval.

The same methodology can be applied to the random freestream perturbation case. Figures 4.9 and 4.10 show the estimator performance when the freestream velocity is randomly perturbed (cutoff frequency is 2). Different from the deceleration case, the perturbations are already present at the beginning of the simulation, and the estimator must be able to filter out the effects of the unknown initial condition while tracking the changes in the freestream. This fact accounts for the longer transient that the estimator experiences. Again, the explicit stochastic forcing plays a fundamental role in keeping the estimator responsive to changes in the freestream velocity at late estimation times.

4.2 Uncertain Reynolds Number

As it was pointed out in the introduction, one of the most common drawbacks of reduced-order models is its fragility to initial conditions and parameters such as the Reynolds number. This characteristics limits the applicability of such models outside of its design envelope. A popular approach to this problem is to create a library of models that spans the parameter space and interpolate between them as needed[97]. Because it uses a full-rank model, the EnKF inherits its robustness to initial conditions, and gives us freedom to choose the initial ensemble as we see fit. As seen in section 3.1.2, the initialization of the ensemble plays an important role
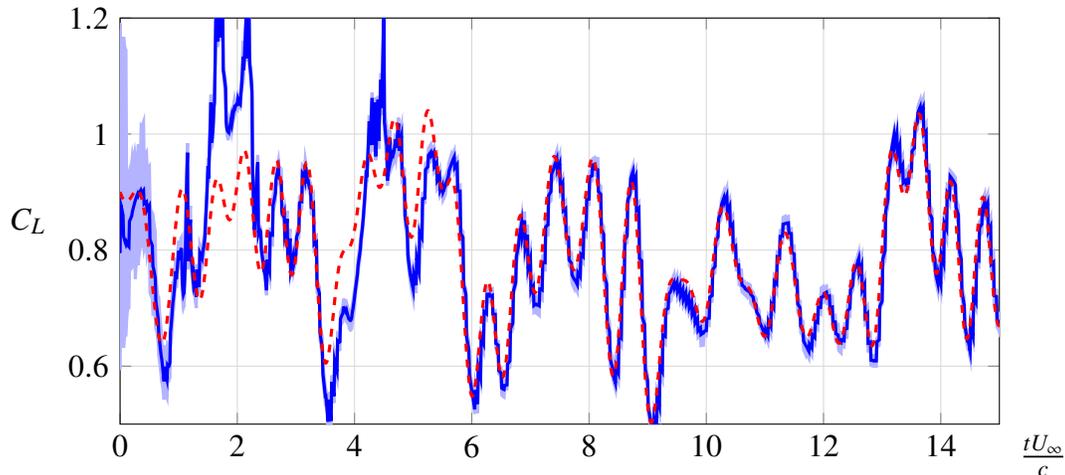
Figure 4.9: Estimated lift coefficient for an airfoil subjected to random freestream perturbation (reduced cutoff frequency is two) using the RTPS covariance inflation scheme ($\theta = 0.9$) and $R = 10^{-6}I_p$. The dashed line (- - -), the solid line (——), and the shaded area (▨) represent, respectively, the reference solution, the EnKF estimate, and the ensemble min-max envelope.
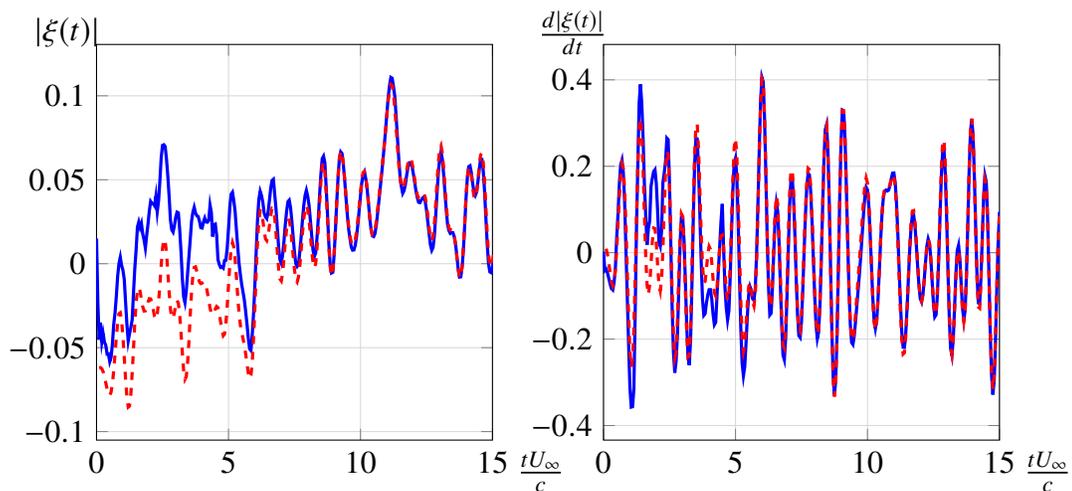
(a) Free stream velocity perturbation.

(b) Free stream velocity perturbation derivative.

Figure 4.10: Joint parameter/state estimation for an airfoil subjected to random freestream perturbations (reduced cutoff frequency is two). The measurement noise level is set to $R = 10^{-6}I_p$ and the RTPS inflation level is set to $\theta = 0.90$. The dashed line (- - -) corresponds to the actual perturbation, and the solid line (——) represent the estimated values with the explicit stochastic forcing in the perturbation derivative dynamics (magnitude 10^{-8}).

in the transient behavior of the filter. If uncertain, any relevant parameter can be estimated on the fly by adopting the augmented state methodology.

In this section, we seek to estimate the Reynolds number corresponding to the reference solution. The state vector is augmented with the estimate for Re, whose dynamics is represented by a persistent model with a lower-bound saturation.

$$\hat{z}_k = \begin{bmatrix} \hat{x}_k \\ \hat{Re}_k \end{bmatrix} = \begin{bmatrix} f(x_{k-1}) \\ \max(Re_{k-1}, 5) \end{bmatrix} + \tilde{\mu}_k = \tilde{f}(z_{k-1}) + \tilde{\mu}_k , \tag{4.6}$$

where $\tilde{\mu}_k \sim N(0, \tilde{Q})$ is the augmented noise vector whose covariance matrix is given by

$$\tilde{Q}_k = \begin{bmatrix} Q & 0 \\ 0 & \sigma \end{bmatrix} . \tag{4.7}$$

Figure 4.11 shows the evolution of the estimate for Re for the flow past a flat plate with pressure measurements. The spatial discretization for the estimator is identical to the reference run, but the parameter initial ensemble is sampled from $Re_0 \sim N(150, 75)$ with any sample below 5 being rejected. Similar to what happened with the freestream velocity estimation, a transient behavior corresponding to the estimation of the vorticity field is observed for early times. After just 2 convective time units, the Re estimate converges to its reference value. In these results, we set $\sigma = 0$, as the tracked parameter is expected to be constant and the only uncertainty is associated with its initial value.
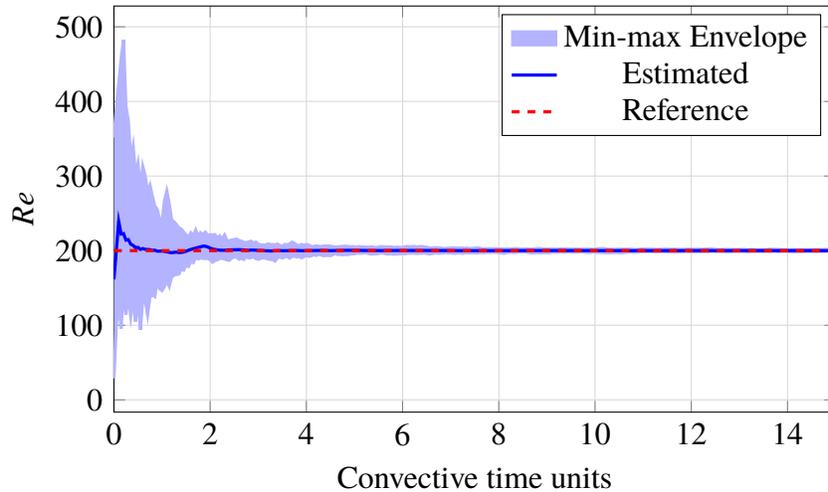
Figure 4.11: Estimated Re for the flow past an inclined flat plate using pressure measuremerts and the RTPS covariance inflation scheme ($\theta = 0.9$ and $R = 10^{-4} I_p$).

4.3   Summary

The results of chapter 3 showed that the EnKF methodology was capable of efficiently estimating the initial condition (represented by the phase of the limit cycle) of a high-dimensional fluid system using limited measurements. In this chapter, we investigated how the augmentation approach can be used to simultaneously estimate model parameters (Reynolds number) and boundary conditions (freestream perturbations). Process noise was shown to play a fundamental role in controlling the estimator long-term receptivity to new events.

In both cases, however, we had prior knowledge that allowed us to create a parametric model to address the source of modeling errors directly. It may be the case, however, that the exact cause of the observed errors are unknown, or that the cost involved in producing an exact cure is prohibitive. In the next chapter, we will focus on modeling the bias (effect) introduced by the model, rather than correcting it (source).

Chapter 5

ESTIMATION IN THE PRESENCE OF RESOLUTION ERROR

If the source of the modeling errors is unknown, or the incremental cost involved in directly fixing the model would render it prohibitive, the presence of these modeling errors can be taken into account by treating them stochastically as process and measurement noise. When these errors have nonzero mean, the bias introduced by the forecast and observation functions can lead the estimator to converge to the wrong solution or become unstable. In the context of ensemble methods when the ensemble size is much smaller than the number of degrees of freedom of the forecast model, biased models can be particularly harmful. Since the corrections applied to the forecast state lie in the low-dimensional subspace spanned by the prior ensemble perturbations, large dynamic bias can render the true state of the system unreachable.

Figure 5.1 highlights two of the main sources of estimation bias. A first source is related to the optimality of the filtering scheme when nonlinear observation functions are present. This source of error is discussed in sections 1.5.6 and 3.3.2. The second source of error is related to the model predictive accuracy. Poor resolution or incorrect representation of the underlying physics can render the model a biased representation of reality. Friedland [98] was one of the first to propose a direct treatment of the forecast error. He proposed a two-stage sequential estimator in which state and bias vector were treated independently, termed the Separate-bias Kalman Filters (SepKF). Later, Dee and Da Silva [99] built upon previous work to derive a rigorous method to independently estimate and sequentially correct for forecast bias. Drecourt, Madsen, and Rosbjerg [100] compared this method to the colored-noise Kalman Filter (ColKF), in which the state vector is augmented to account for noise processes modeled by autoregressive models. These works, however, all assume that the observation model is unbiased.

In this chapter, we develop a bias-aware estimator that allows us to use an ensemble of coarse-grid simulations to robustly estimate both the state and the output of a fine-grid reference run.
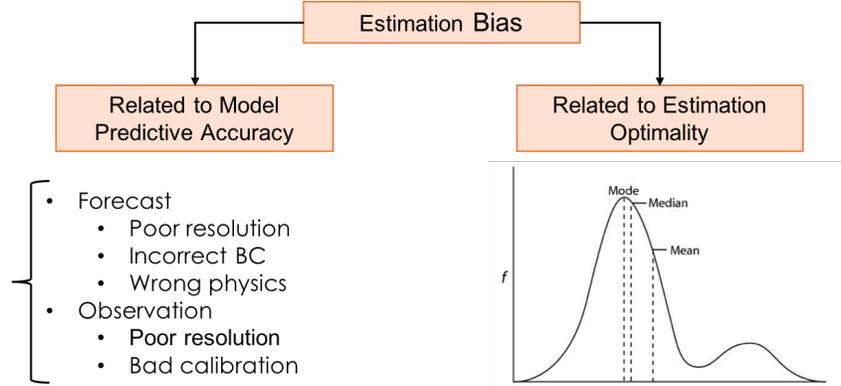
Figure 5.1: Sources of bias in an Kalman-based estimator.

5.1 Resolution Error as a Source of Bias

Let $\tilde{f}(x)$ be the exact transition function of the Navier-Stokes equations, which maps the infinite-dimensional solution $\tilde{x}_{k-1}$ at time t_{k-1} to the solution $\tilde{x}_k$ at time t_k,

$$\tilde{x}_k = \tilde{f}(\tilde{x}_{k-1}) \,, \tag{5.1a}$$

$$y_k = \tilde{h}(\tilde{x}_k) + \epsilon_k^m \,, \tag{5.1b}$$

where $\epsilon_k^m \sim N(0, \tilde{R}_k)$ is a p-dimensional random error vector associated with the measurement methodology that is independent of the state and uncorrelated in time. Since both the state $\tilde{x}$ and the operators $\tilde{f}$ and $\tilde{h}$ are unattainable for practical purposes, we introduce a finite-dimensional approximation for the model and state. We notate the finite-dimensional approximations with the same symbols but without the tilde.

Following Cohn[74], we define a projection operator Π that maps the the true state $\tilde{x}_k$ onto its finite-dimensional representation $x_k = \Pi \tilde{x}_k$. The propagation of x_k can be represented as

$$x_k = f(x_{k-1}) + \delta_{k-1} \,, \tag{5.2}$$

where

$$\delta_k = \Pi \tilde{f}(\tilde{x}_k) - f(\Pi \tilde{x}_k) \,. \tag{5.3}$$

The forcing term δ_k represents the *model error*, and gathers errors from different sources: discretization error, inaccurate boundary conditions, uncertain forcing, and so on.

Analogously, since the continuous state is never available for practical purposes, a discrete version of the observation operator $h(\cdot)$ needs to be introduced:

$$y_k = h_k(x_k) + \epsilon_k^m + \epsilon_k^r \,, \tag{5.4}$$

where

$$\begin{aligned} \epsilon_k^r &= \tilde{h}(\tilde{x}_k) - h(\Pi\tilde{x}_k) \\ &= (\tilde{h}(\tilde{x}_k) - \tilde{h}(\Pi\tilde{x}_k)) + (\tilde{h}(\Pi\tilde{x}_k) - h(\Pi\tilde{x}_k)) \end{aligned} \tag{5.5}$$

is the *error of representativeness* [74]. This error can be further split into two contributions: a first term that represents the effect of the exact operator on the unresolved scales (aliasing), and a second term that represents the discretization error of the operator itself.

5.2   Low-rank Representation of the Bias

Because δ_k depends not only on the state but also on the continuous operator $\tilde{f}$, its value is unknowable from a deterministic point of view. Therefore, the most common approach is to represent this error as a stochastic perturbation with known bias and covariance.

Thus, we use

$$\delta_k = \Gamma_x \xi_k + \mu_k \,, \tag{5.6}$$

where Γ_x is a low-rank representation of the low-frequency behavior of the bias and represents the available deterministic knowledge about the model error. The second term ($\mu_k \sim N(0, Q_k)$) represents the high-frequency portion of the bias, and is modeled as white. In meteorology, Dee [101] and Cohn and Parrish [102] have proposed models for Q_k whose corresponding parameters can be tuned on-line. These models often rely on a low-order representation of these matrices (a popular choice is to use the slow modes of the forecast model).

A complete representation of δ_k requires an impractical amount of data and operations. Since the dimension p of the measurement vector obtained at any given time is customarily smaller than the number n of degrees of freedom of the forecast model, it is not possible to estimate all $n(n + 1)/2$ degrees of freedom of the covariance matrix associated to δ_k in real time, regardless of the estimation procedure used [101].

The measurement resolution error can also be split in two terms:

$$\epsilon_k^r = \Gamma_y \eta_k + \nu_k \, , \tag{5.7}$$

where Γ_y is a low-rank representation of the low-frequency behavior of the bias, and $\nu_k \sim N(0, R^r)$ represents the high-frequency content of the error in the measurement error. The measurement accuracy error ϵ_k^m is merged to ν_k, which represent the zero-mean high-frequency portion of the observation error, which is now represented as zero-mean Gaussian process with a combined covariance matrix R.

It is to be expected that ξ_k and η_k are auto-correlated in time. As a first option, autoregressive models (AR)[1] are the simplest way of representing this feature in the discrete-time framework. Noise sequences obeying AR models are often referred to as colored-noise sequences [100, 103]. If N_x and N_y are the orders for the autoregressive models for the state and observation bias, respectively, then

$$\xi_k = \bar{\xi} + \sum_{i=1}^{N_x} \Phi_i^x (\xi_{k-i} - \bar{\xi}) + \gamma_{x,k} \tag{5.8}$$

$$\eta_k = \bar{\eta} + \sum_{i=1}^{N_y} \Phi_i^y (\eta_{k-i} - \bar{\eta}) + \gamma_{y,k} \, , \tag{5.9}$$

where $\gamma_{x,k} \sim N(0, Q^b)$ and $\gamma_{y,k} \sim N(0, R^b)$, and Φ_i^x and Φ_i^y should be determined based on prior knowledge about the system under study. For example, if the Markov condition holds ($N_x = N_y = 1$) and these error terms are supposed to vary in a much slower time scale than the state dynamics, one can choose to use a persistent model and set $\Phi^x = \Phi^y = 1$.

A second possibility that can be more suitable for periodic flows is to represent the bias as a sum of harmonics of the system's characteristic frequency. In this framework, the columns of Γ_x and Γ_y and the bias coefficient ξ_k and η_k are represented by complex vectors, and the bias dynamics are then given by

$$\delta_k = \Re(\Gamma_x \xi_k) + \mu_k \tag{5.10}$$

$$\epsilon_k^r = \Re(\Gamma_y \eta_k) + \nu_k \, , \tag{5.11}$$

and

$$\xi_k = \exp(\Lambda_x \Delta t) \xi_{k-1} + \gamma_{x,k} \tag{5.12}$$

$$\eta_k = \exp(\Lambda_y \Delta t) \eta_{k-1} + \gamma_{y,k} \, , \tag{5.13}$$

---

[1] AR models differ from moving-average (MA) models by the fact that in the former the weights are applied to the previous states, while in the latter the weights are applied to the input (noise).

where Λ_x and Λ_y are diagonal matrices whose entries are also complex numbers. The real part of these entries represents the mode's growth/decay rate, which is expected to be zero in a purely periodic flow, and the corresponding imaginary part is related to the mode's oscillatory frequency. Note that the noise terms γ_x and γ_y are now complex random sequences with zero-mean Gaussian-distributed magnitude, but uniformly-distributed phase.

In order to take advantage of an eventual nonzero cross-correlation between the bias parameters ξ and η and the state x in the filtering process, these variables are gathered in an augmented state vector ($z = \begin{bmatrix} x^T \xi^T \eta^T \end{bmatrix}^T$) that will be estimated using the algorithm proposed in the next section.

5.3 A Bias-aware Ensemble Kalman Filter

Having the standard ensemble Kalman filter presented in section 1.5 as a starting point, bias-awareness can be achieved using the augmentation approach: the parameters corresponding to the low-rank bias models developed in the previous section are appended to the original state vector.

The modifications to the two filtering steps can be summarized by:

Forecast Step: The state of each ensemble member (here denoted by the superscript (j)) at the next time step is forecast using the (possibly nonlinear) dynamic model[2]:

$$
\hat{z}_k^{(j)} = \begin{bmatrix} \hat{x}_k^{(j)} \\ \hat{\xi}_k^{(j)} \\ \hat{\eta}_k^{(j)} \end{bmatrix} = f^\dagger(z_{k-1}^{(j)}) + \mu_k^{\dagger(j)}
$$

$$
= \begin{bmatrix} f(x_{k-1}^{(j)}) \\ \bar{\xi} + \Phi_x(\xi_{k-1}^{(j)} - \bar{\xi}) \\ \bar{\eta} + \Phi_y(\eta_{k-1}^{(j)} - \bar{\eta}) \end{bmatrix} + \begin{bmatrix} \mu_k^{(j)} \\ \gamma_{x,k}^{(j)} \\ \gamma_{y,k}^{(j)} \end{bmatrix} . \tag{5.14}
$$

Assuming $n_s \ll n$ and $n_o \ll n$, the additional cost associated with the bias dynamics should be negligible.

---

[2]Here the bias dynamics is represented by the AR1 model, but the algorithm can be easily adapted to use higher-order auto-regressive models or the DMD-based periodic model.

Thus, the prior statistics can be evaluated as

$$\hat{Z}_k = \begin{bmatrix} \hat{z}_k^{(1)} & \hat{z}_k^{(2)} & \cdots & \hat{z}_k^{(q)} \end{bmatrix} \tag{5.15a}$$

$$\bar{z}_k = \frac{1}{q} \hat{Z}_k \mathbb{1} \tag{5.15b}$$

$$\hat{C}_k^{zz} = \frac{1}{\sqrt{q-1}} \sum_{i=1}^{q} \left(\hat{z}_k^{(i)} - \bar{z}_k \right) \left(\hat{z}_k^{(i)} - \bar{z}_k \right)^T , \tag{5.15c}$$

which can also be expressed in terms of the scaled ensemble perturbation matrix

$$\hat{A}_k = \frac{1}{\sqrt{q-1}} \hat{Z}_k \left(I - \bar{z}_k \mathbb{1}^T \right) \tag{5.16a}$$

$$\hat{C}_k^{zz} = \hat{A}^{(j)} \left(\hat{A}^{(j)} \right)^T . \tag{5.16b}$$

Analysis Step: Taking after what was shown in section 1.5.1, we frame the analysis step as an optimization problem. For each of the ensemble members, provided that both the prior and the measurement distributions are Gaussian, the mode (maximum-likelihood estimate) of the posterior distribution corresponds to the minimizer of the cost function:

$$J(z) = \frac{1}{2\alpha} \left\| G^{-1}z - \hat{z}_k^{(j)} \right\|_{\hat{C}_k^{zz}}^2 + \frac{1}{2} \left\| y_k - h(x) - \Gamma_y \eta - v_k^{(j)} \right\|_R^2 , \tag{5.17}$$

while restricting $[x - \Gamma_x \xi \quad \xi \quad \eta]^T = G^{-1}z$ to the affine subset generated by the prior estimate of the state $\hat{z}_k^{(j)}$ and the subspace spanned by the scaled perturbation matrix $\hat{A}_k$, i.e.,

$$z_k = \operatorname*{argmin}_{z \in G(\hat{z} + \hat{A}_k)} J(z) , \tag{5.18}$$

where

$$G = \begin{bmatrix} I & \Gamma_x & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} . \tag{5.19}$$

The first term in the cost function acts as an regularization term by penalizing the distance of the state (before bias correction) to the prior estimate, and the second term penalizes the data mismatch between the observed measurement y_k and the ones predicted by the observation model (after bias correction). We also include the parameter α representing the magnitude of the multiplicative covariance inflation. In the simulation that led to the results shown in this chapter, the RTPS scheme was used (see section 1.5.4 for details).

5.3.1 The Linear Observation Function Case

When the observation function is linear, i.e. $h(x) = Hx$, the observation equation can be written as

$$\hat{y}_k^{(j)} = \tilde{H}G\hat{z}_k^{(j)} + v_k^{(j)} , \qquad (5.20)$$

where

$$\tilde{H} = \begin{bmatrix} H & 0 & \Gamma_y \end{bmatrix} . \qquad (5.21)$$

The restriction on the optimization space is enforced by means of a change of variables

$$z = G\left(\hat{z}_k^{(j)} + \sqrt{\alpha}\hat{A}_k v\right) , \qquad (5.22)$$

where $v \in \mathbb{R}^q$ is the correction coefficient vector.

The analysis step objective is then defined as finding

$$v_k = \underset{v \in \mathbb{R}^q}{\operatorname{argmin}} \, J(v) , \qquad (5.23)$$

where

$$J(v) = \frac{1}{2} \|v\|^2 + \frac{1}{2} \left\| y_k - \tilde{H}G\left(\hat{z}_k^{(j)} + \sqrt{\alpha}\hat{A}_k v\right) - v_k^{(j)} \right\|_R^2 . \qquad (5.24)$$

Since $J(v)$ is quadratic in v, the solution is unique and corresponds to the root of

$$DJ(v) = v - \sqrt{\alpha}\left(\tilde{H}G\hat{A}_k\right)^T R^{-1}\left(y_k - \tilde{H}G\hat{z}_k^{(j)} - \sqrt{\alpha}\tilde{H}G\hat{A}_k v - v_k^{(j)}\right) = 0 , \qquad (5.25)$$

which is given by

$$v_k^{(j)} = \sqrt{\alpha}\left[I + \alpha\left(\tilde{H}G\hat{A}_k\right)^T R^{-1}\left(\tilde{H}G\hat{A}_k\right)\right]^{-1}(\tilde{H}G\hat{A}_k)^T R^{-1}\left(y_k - \tilde{H}G\hat{z}_k^{(j)} - v_k^{(j)}\right)$$

$$(5.26a)$$

$$= \sqrt{\alpha}\left(\tilde{H}G\hat{A}_k\right)^T \left[R + \alpha\left(\tilde{H}G\hat{A}_k\right)\left(\tilde{H}G\hat{A}_k\right)^T\right]^{-1}\left(y_k - \tilde{H}G\hat{z}_k^{(j)} - v_k^{(j)}\right) ,$$

$$(5.26b)$$

where we have used the Woodbury matrix identity to obtain the second line.

When the inversion is performed in the ensemble space (Eq. 5.26a), the final solution is obtained by projecting these coefficients back to the state space:

$$z_k^{(j)} = G\hat{z}_k^{(j)} + \alpha G\hat{A}_k \left[I + \alpha \left(\tilde{H}G\hat{A}_k \right)^T R^{-1} \left(\tilde{H}G\hat{A}_k \right) \right]^{-1} (\tilde{H}G\hat{A}_k)^T R^{-1} \times$$
$$\left(y_k - \tilde{H}G\hat{z}_k^{(j)} - v_k^{(j)} \right) \tag{5.27a}$$

$$= G\hat{z}_k^{(j)} + \alpha G\hat{A}_k \left(\tilde{H}G\hat{A}_k \right)^T \left[R + \alpha \left(\tilde{H}G\hat{A}_k \right) \left(\tilde{H}G\hat{A}_k \right)^T \right]^{-1} \times$$
$$\left(y_k - \tilde{H}G\hat{z}_k^{(j)} - v_k^{(j)} \right) . \tag{5.27b}$$

Alternatively, when the inversion is performed in the measurement space, instead of solving for v_k, the representers formulation[57] is used:

$$\left[R + \alpha \left(\tilde{H}G\hat{A}_k \right) \left(\tilde{H}G\hat{A}_k \right)^T \right] b_k^{(j)} = y_k - \tilde{H}G\hat{z}_k^{(j)} - v_k^{(j)} \tag{5.28a}$$

$$z_k^{(j)} = G\hat{z}_k^{(j)} + \alpha G\hat{A}_k \left(\tilde{H}G\hat{A}_k \right)^T b_k^{(j)} . \tag{5.28b}$$

5.3.2 The Nonlinear Observation Function Case

In section 1.5.6, we presented approximate schemes that generalizes the EnKF to the case when $h(x)$ is nonlinear. In the following sections, we follow the same steps to obtain the modified equations corresponding to the bias-ware filter.

Implicit Linearization   This approach can be understood as an approximation to an extended Kalman filter in which the observation function is linearized about the ensemble mean. Since the linearized operator is never explicitly computed, we use the expression *implicit linearization* to refer to this scheme. We again start by augmenting the state vector with the predicted measurements, and redefining the

observation function to select the last variable of the augmented state vector.

$$
\hat{w}_k = \begin{bmatrix} \hat{x}_k \\ \hat{\xi}_k \\ \hat{\eta}_k \\ \hat{y}_k \end{bmatrix} = f^\dagger(w_{k-1}) + \mu_k^\dagger
$$

$$
= \begin{bmatrix} f(x_{k-1}) \\ \Phi_x \xi_{k-1} + (I - \Phi_x)\bar{\xi} \\ \Phi_y \eta_{k-1} + (I - \Phi_y)\bar{\eta} \\ h(f(x_{k-1}) + \Gamma_x \xi_k) + \Gamma_y \eta_k \end{bmatrix} + \begin{bmatrix} \mu_k \\ \gamma_{x,k} \\ \gamma_{y,k} \\ \nu_k \end{bmatrix} \tag{5.29}
$$

$$
\hat{y}_k = \begin{bmatrix} 0 & 0 & 0 & I \end{bmatrix} \hat{w}_k = L\hat{w}_k . \tag{5.30}
$$

The associated cost function is

$$
J(w) = \frac{1}{2\alpha} \left\| \tilde{G}^{-1} w - \hat{w}_k \right\|^2_{\hat{C}_k^{ww}} + \frac{1}{2} \left\| y_k - Lw \right\|^2_R , \tag{5.31}
$$

where

$$
\tilde{G} = \begin{bmatrix} G & 0 \\ 0 & I \end{bmatrix} \tag{5.32}
$$

and

$$
\hat{C}_k^{ww} = \begin{bmatrix} \hat{C}_k^{zz} & \hat{C}_k^{zy} \\ \left(\hat{C}_k^{zy}\right)^T & \hat{C}_k^{yy} \end{bmatrix} = \begin{bmatrix} \hat{C}_k^{xx} & \hat{C}_k^{x\xi} & \hat{C}_k^{x\eta} & \hat{C}_k^{xy} \\ \left(\hat{C}_k^{x\xi}\right)^T & \hat{C}_k^{\xi\xi} & \hat{C}_k^{\xi\eta} & \hat{C}_k^{\xi y} \\ \left(\hat{C}_k^{x\eta}\right)^T & \left(\hat{C}_k^{\xi\eta}\right)^T & \hat{C}_k^{\eta\eta} & \hat{C}_k^{\eta y} \\ \left(\hat{C}_k^{xy}\right)^T & \left(\hat{C}_k^{\xi y}\right)^T & \left(\hat{C}_k^{\eta y}\right)^T & \hat{C}_k^{yy} \end{bmatrix} \tag{5.33a}
$$

$$
= \frac{1}{q-1} \sum_{j=1}^{q} (\hat{w}_k^j - \bar{w}_k)(\hat{w}_k^j - \bar{w}_k)^T \tag{5.33b}
$$

$$
= \hat{A}_k^w \left(\hat{A}_k^w\right)^T . \tag{5.33c}
$$

Using a change of variables similar to the one proposed in Eq. 5.22, Eq. 5.31 can be rewritten as

$$
w = \tilde{G}\left(\hat{w}_k^{(j)} + \sqrt{\alpha}\hat{A}_k^w v\right) \tag{5.34}
$$

$$
J(v) = \frac{1}{2} \|v\|^2 + \frac{1}{2} \left\| y_k - L\tilde{G}(\hat{z}_k^{(j)} + \sqrt{\alpha}\hat{A}_k v) \right\|^2_R . \tag{5.35}
$$

Since this function is quadratic in v, the minimizer is given by

$$v_k^{(j)} = \sqrt{\alpha} \left[I + \alpha \left(L\tilde{G}\hat{A}_k^w \right)^T R^{-1} \left(L\tilde{G}\hat{A}_k^w \right) \right]^{-1} (L\tilde{G}\hat{A}_k^w)^T R^{-1} \left(y_k - L\tilde{G}\hat{w}_k^{(j)} \right)$$

(5.36a)

$$= \sqrt{\alpha} \left(L\tilde{G}\hat{A}_k^w \right)^T \left[R + \alpha \left(L\tilde{G}\hat{A}_k^w \right) \left(L\tilde{G}\hat{A}_k^w \right)^T \right]^{-1} \left(y_k - L\tilde{G}\hat{w}_k^{(j)} \right) , \quad (5.36b)$$

where

$$L\tilde{G}\hat{w}_k^{(j)} = \hat{y}_k^{(j)} = h\left(f(\hat{x}_{k-1}^{(j)}) + \Gamma_x \hat{\xi}_k^{(j)} \right) + \Gamma_y \hat{\eta}_k^{(j)} + v_k^{(j)} = h^\dagger \left(G\hat{z}_k^{(j)} \right) + v_k^{(j)} \quad (5.37)$$

$$\left(L\tilde{G}\hat{A}_k^w \right) \left(L\tilde{G}\hat{A}_k^w \right)^T = \hat{C}_k^{yy} . \quad (5.38)$$

The posterior solution is then given by

$$w_k^{(j)} = \tilde{G}\hat{w}_k^{(j)} + \alpha\tilde{G}\hat{A}_k^w \left[I + \alpha \left(L\tilde{G}\hat{A}_k^w \right)^T R^{-1} \left(L\tilde{G}\hat{A}_k^w \right) \right]^{-1} (L\tilde{G}\hat{A}_k^w)^T \tilde{G} R^{-1} \times$$

$$\left(y_k - L\tilde{G}\hat{w}_k^{(j)} \right) \quad (5.39a)$$

$$= \tilde{G}\hat{w}_k^{(j)} + \alpha\tilde{G}\hat{A}_k^w \left(L\tilde{G}\hat{A}_k^w \right)^T \left[R + \alpha \left(L\tilde{G}\hat{A}_k^w \right) \left(L\tilde{G}\hat{A}_k^w \right)^T \right]^{-1} \times$$

$$\left(y_k - L\tilde{G}\hat{w}_k^{(j)} \right) . \quad (5.39b)$$

Iterative optimization In this approach, the posterior mean is approximated by the posterior mode, which is obtained by directly minimizing

$$J(v) = \frac{1}{2}\|v\|^2 + \frac{1}{2} \left\| y_k - h^\dagger \left(G(\hat{z}_k^{(j)} + \sqrt{\alpha}\hat{A}_k v) \right) - v_k^{(j)} \right\|_R^2 \quad (5.40)$$

using an iterative method to find the root of

$$DJ(v) = v - \sqrt{\alpha} \left[\frac{\partial h^\dagger}{\partial z} \left(G(\hat{z}_k^{(j)} + \sqrt{\alpha}\hat{A}_k v) \right) G\tilde{A}_k \right]^T R^{-1} \times$$

$$\left(y_k - h^\dagger \left(G(\hat{z}_k^{(j)} + \sqrt{\alpha}\hat{A}_k v) \right) - v_k^{(j)} \right) = 0 . \quad (5.41)$$

Employing the damped Gauss-Newton scheme presented in section 1.5.6, one obtains

$$v^{i+1} = (1 - \beta)v^i + \beta \left(I + B(v^i)^T R^{-1} B(v^i) \right)^{-1} B(v^i)^T R^{-1} \times$$

$$\left[y_k - h^\dagger \left(G(\hat{z}_k^{(j)} + \sqrt{\alpha}\hat{A}_k v^i) \right) + B(v^i)v^i - v_k^{(j)} \right] \quad (5.42a)$$

$$= (1 - \beta)v^i + \beta B(v^i)^T \left(R + B(v^i)B(v^i)^T \right)^{-1} \times$$

$$\left[y_k - h^\dagger \left(G(\hat{z}_k^{(j)} + \sqrt{\alpha}\hat{A}_k v^i) \right) + B(v^i)v^i - v_k^{(j)} \right] , \quad (5.42b)$$

where

$$B(v) = \sqrt{\alpha} \frac{\partial h^{\dagger}}{\partial z} \left(G(\hat{z}_k^{(j)} + \sqrt{\alpha} \hat{A}_k v) \right) G \hat{A}_k$$
$$= \sqrt{\alpha} \left[\frac{\partial h}{\partial x} \left([I\ 0\ 0] G(\hat{z}_k^{(j)} + \sqrt{\alpha} \hat{A}_k v) \right) \quad 0 \quad \Gamma_y \right] G \hat{A}_k \tag{5.43}$$

and β is the first value in the sequence $1, 1/2, 1/4, \ldots, 2^{-n}$ for which the Armijo-Goldstein[72] step principle (Eq. 1.88) is satisfied. The iterative process ends when $\|v^{i+1} - v^i\|_{\infty} \leq \epsilon_1$, $J(v^i) - J(v^{i+1}) \leq \epsilon_2 J(v^i)$ or the pre-set maximum number of iterations is exceeded I_{max}. In this work, $\epsilon_1 = 10^{-6}$, $\epsilon_2 = 10^{-4}$ and $I_{max} = 10$.

5.4 Identification of the Grid-resolution Model Error

The numerical error introduced by the different resolution levels is the source of the bias that we will be interested in tracking. State statistics are estimated from a dataset of the base solutions spanning a sufficiently long time window. Bias statistics can then be estimated using the definitions presented in section 5.1:

$$\Delta = \left[\Pi x_2^f - f\left(\Pi x_1^f \right) \quad \cdots \quad \Pi x_n^f - f\left(\Pi x_{n-1}^f \right) \right] \tag{5.44a}$$

$$\bar{\delta} = \frac{1}{n-1} \Delta \mathbb{1} \tag{5.44b}$$

$$E = \left[\Pi h^f \left(x_1^f \right) - h^c \left(\Pi x_1^f \right) \quad \cdots \quad \Pi h^f \left(x_n^f \right) - h^c \left(\Pi x_{n-1}^f \right) \right] \tag{5.44c}$$

$$\bar{\epsilon}^r = \frac{1}{n-1} E \mathbb{1} , \tag{5.44d}$$

where the superscripts f and c correspond, respectively, to the fine and coarse meshes, and Π is the interpolation operator between the fine and coarse meshes.

Figures 5.2 and 5.3 show the temporal mean of the bias fields of both the dynamics (δ) and observation model (ϵ^r) between the $Re_\Delta = 4$ and $Re_\Delta = 1$ resolution levels for both the airfoil and the flat plate case. In both cases, the mean bias in the dynamics seems to concentrate near the body where the error introduced by the immersed boundary dominates. Regarding the observation model, while in the flat plate case most of the bias seems to be restricted to the leading and trailing edges region, the airfoil case experiences a much more pronounced observation bias in the entire surface. The large bias observed in the normal stresses for this case can be explained by the fact that for a closed body the distribution of normal component of the forces acting on it surface is only defined up to a constant (since a constant normal force acting on an immersed body will have zero resultant).
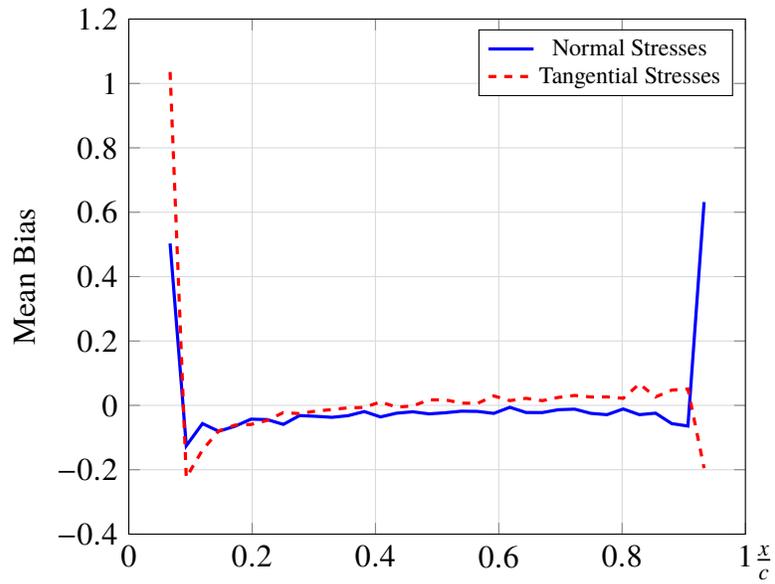
(a) Bias field for the observation model ($\bar{\epsilon}^r$).
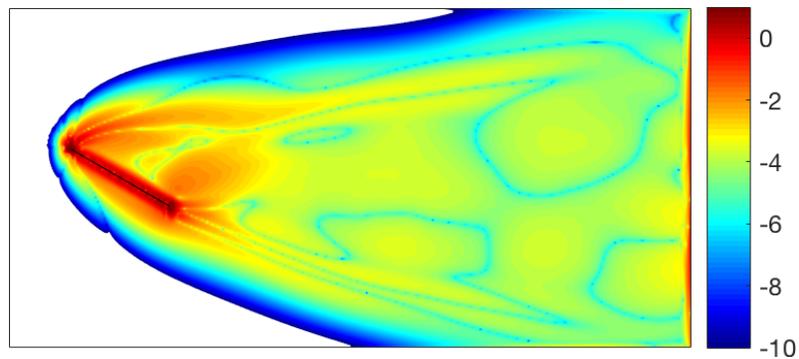
(b) Bias field for the dynamic model ($\bar{\delta}$) in log scale.

Figure 5.2: Temporal average of the bias fields introduced by the resolution error for the flow past an inclined flat plate when comparing the $Re_\Delta = 1$ (200 grid points per chord) simulation to the corresponding $Re_\Delta = 4$ (50 grid points per chord) simulation.
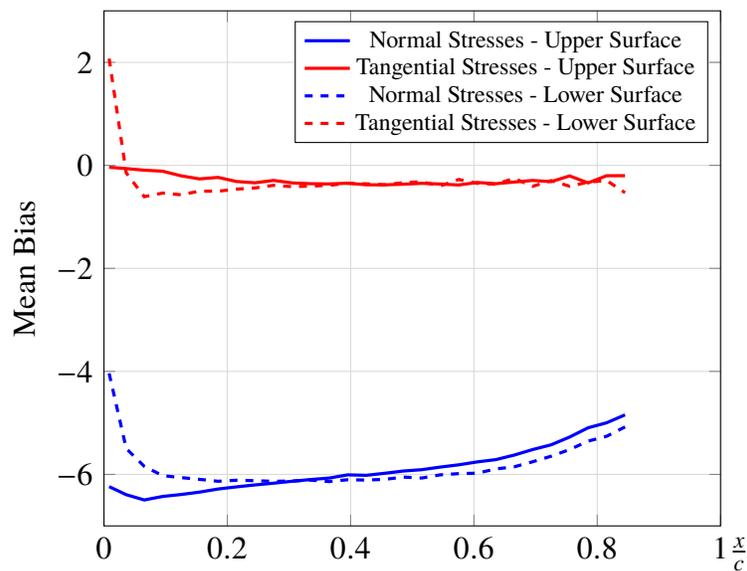
(a) Bias field for the observation model ($\bar{\epsilon}^r$).
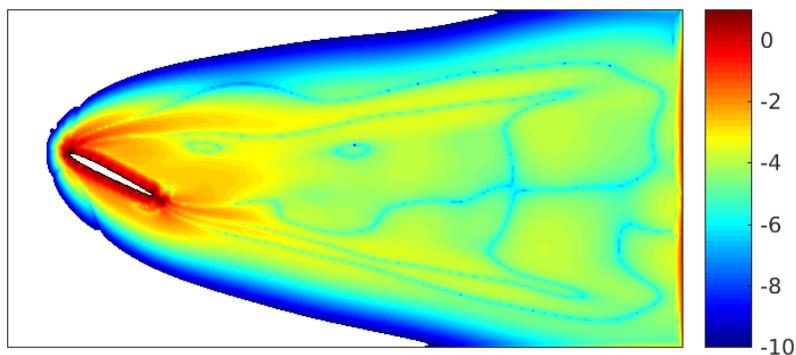
(b) Bias field for the dynamic model ($\bar{\delta}$) in log scale.

Figure 5.3: Temporal average of the bias fields introduced by the resolution error for the flow past an inclined NACA 0009 when comparing the $Re_\Delta = 1$ (200 grid points per chord) simulation to the corresponding $Re_\Delta = 4$ (50 grid points per chord) simulation.

(a) State bias.

(b) Observation (pressure) bias.
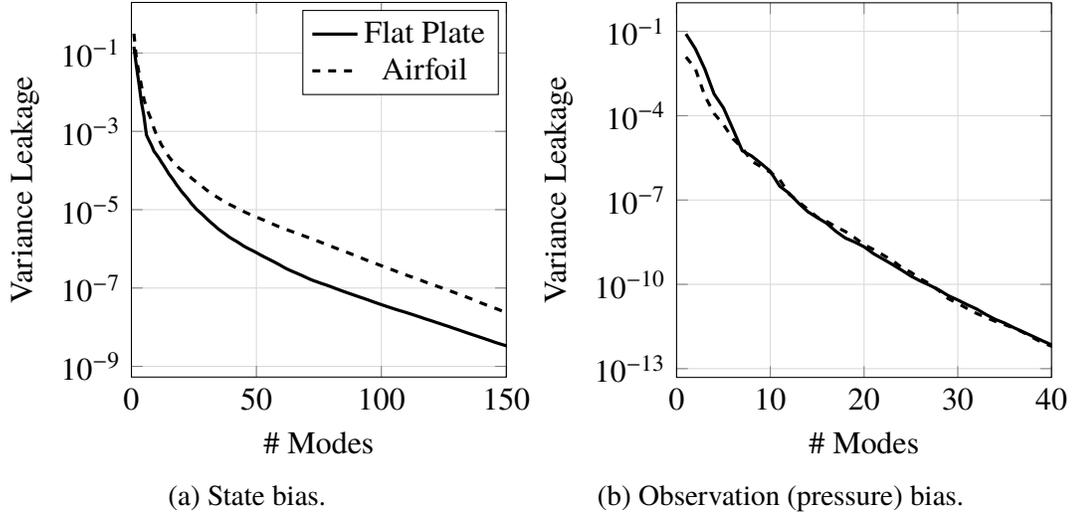
Figure 5.4: Fraction of the bias variance left out by using the corresponding first POD modes for the flow past an inclined flat plate and a NACA 0009 airfoil.

The structure of the corresponding state and observation bias covariance matrices can be analyzed by computing its Proper Orthogonal Decomposition (POD), which can be accomplished by computing the left singular vectors of $\Delta - \bar{\delta}$ and $E - \bar{\epsilon}^r$, respectively. The low-rank bias representation proposed in Sec. 5.2 is justified on the fact that most of the bias variance is restricted to just a few directions in the state space. As Fig. 5.4 indicates that for the flat plate and airfoil, respectively, retaining the first $n_s = 25$ state POD modes and $n_o = 10$ observation POD modes leaves less than 0.01% of the variance to be modeled as white noise.

Therefore, matrices Γ_x, Γ_y, $\bar{\xi}$ and $\bar{\eta}$ in Eqs. 5.8 and 5.9 can be defined as

$$\Gamma_x = \begin{bmatrix} \bar{\delta} & u_1^s & \cdots & u_{n_s}^s \end{bmatrix} \tag{5.45a}$$

$$\Gamma_y = \begin{bmatrix} \bar{\epsilon}^r & u_1^o & \cdots & u_{n_o}^o \end{bmatrix} \tag{5.45b}$$

$$\bar{\xi} = \bar{\eta} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T , \tag{5.45c}$$

where u_i^s and u_i^o are the i-th leading POD modes of of $\Delta - \bar{\delta}$ and $E - \bar{\epsilon}^r$, respectively, normalized by their respective variances. The auto-regressive model parameters are set to $\phi_1^x = \phi_1^y = e^{-\Delta t/\tau}$, where Δt is the time interval between two analysis steps, and τ is a reference decorrelation time, here considered to be the vortex shedding period. Alternatively, one could use a least-squares approach to determine the AR coefficients that best fit the data used to construct the low-rank model. Figure 5.5 shows the prediction error for the best ARn model for each of the columns of Γ_x

(a) State bias.

(b) Observation (pressure) bias.
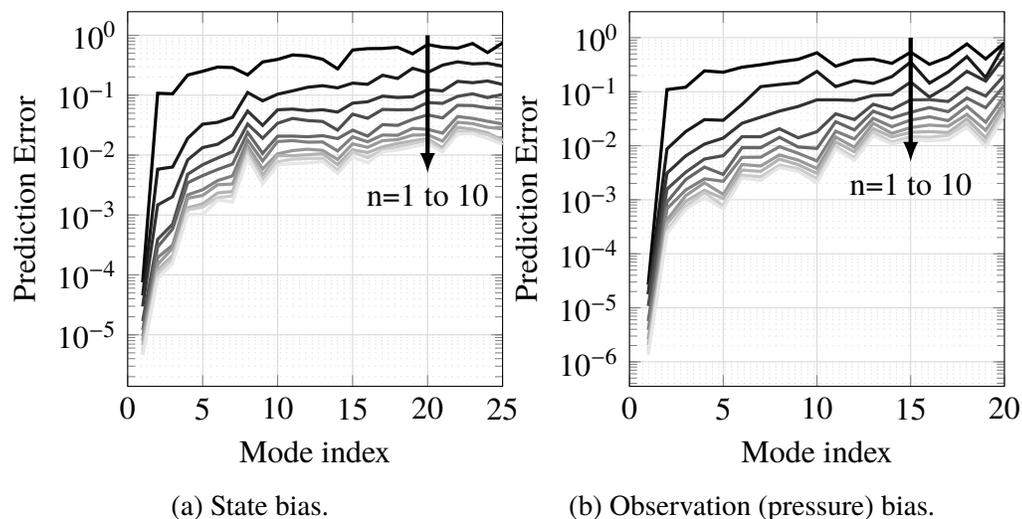
Figure 5.5: Prediction error corresponding to the best (in a least-squares sense) ARn model for each of the POD modes selected to represent the bias. The order of the corresponding auto-regressive model increases from 1 to 10 as the color of the curves changes from dark to light.

and Γ_y. Note that the error for the first mode, the mean, is already low for the AR1 model, since its coefficient is expected to remain constant. The error for the remaining modes, however, decay slowly with increasing model order, indicating that they are more strongly time-correlated. Even though the AR1 coefficients obtained via least-squares differ from the initial guess based on a decorrelation time equal to the vortex shedding period, the differences in performance of the resulting estimator were small[3].

A second approach to characterize the time coherency present in the bias is by using the Dynamic Mode Decomposition (DMD) methodology developed by Schmid [104]. Each of the resulting DMD modes describe a spatial structure that evolves in time with a fixed growth/decay rate and oscillatory frequency. For a periodic phenomenon, the growth/decay rate is expected to be very close to zero, i.e., the Ritz values associated with the DMD modes should lay on top of the unitary circle as it is indeed the case as shown in Fig. 5.6. Sorting the modes by their initial magnitude, the leading modes can be selected to form the matrices Γ_x and Γ_y. Figure 5.7 shows the prediction error of the low-rank model with different numbers of DMD modes when tested in the same data used to generate the DMD modes.

---

[3]Higher-order AR models were also tested as alternatives for representing the dynamics of bias. However, the resulting estimator demonstrated to be unstable for $n > 2$ even though the AR models were verified to be stable themselves.

(a) State bias Ritz values.

(b) Observation bias Ritz values.
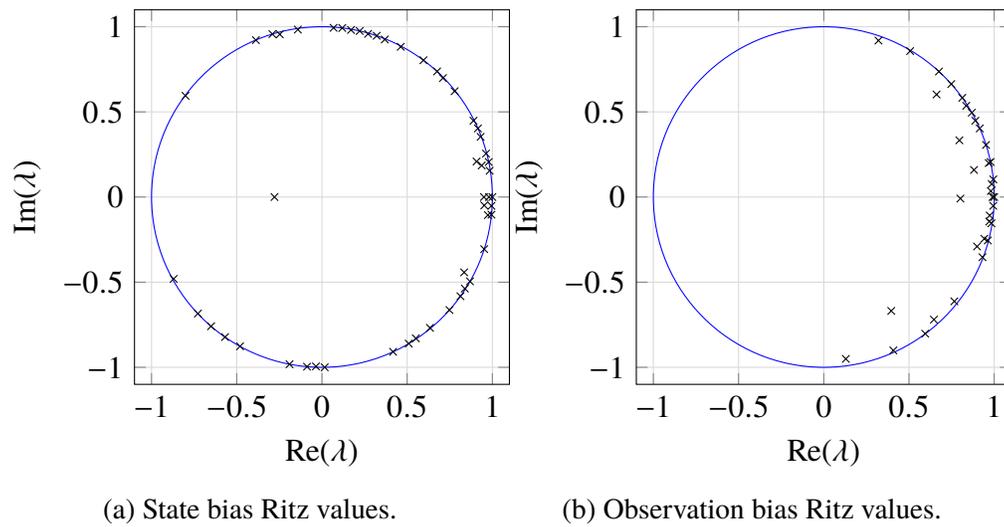
Figure 5.6: Ritz values corresponding to the DMD modes of the bias (when a part of a conjugate pair, only one of them is plotted).

(a) State bias.

(b) Observation (pressure) bias.
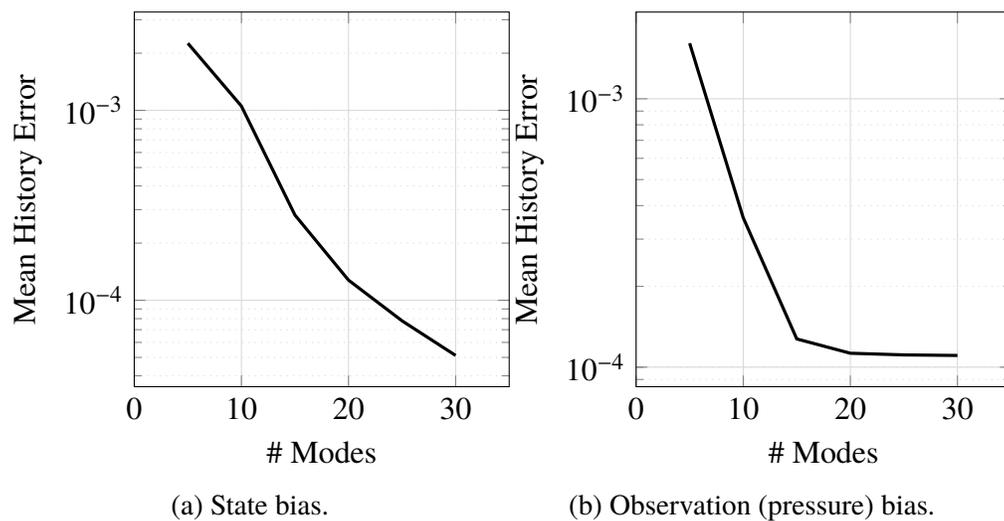
Figure 5.7: Prediction error corresponding to DMD-based low-rank models with different number of DMD modes.
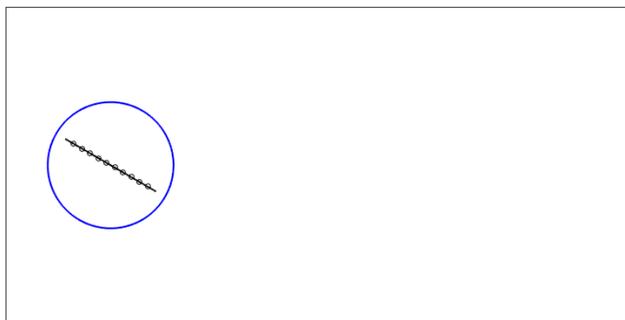
Figure 5.8: Illustration of the different domains used for the evaluation of the state error. The small circles on the surface of the plate represent the locations where the pressure measurements are taken.

5.5 Effect of the Bias

In this section we consider the case when the low-resolution model is used to track the high-resolution flat plate data without an explicit treatment of the model biases. Since most of the bias in the system dynamics is restricted to the region close to the body (see Figs. 5.2 and 5.3), we introduce a second error metric that evaluates the mean-square error of the solution in the region outside the blue circle represented in Fig. 5.8. In this way we can distinguish estimator errors associated with forces on the immersed surface from those associated with the wake dynamics. The measurement error level is set to $R = 10^{-4}$ for all cases henceforward. Process noise can be added to the low-resolution dynamics to inform the estimator of a reduction of the perceived reliability of the dynamic model. Since it is presumed that no information about the bias is available, additive covariance inflation (which is equivalent to process noise in this context) will be sampled from the same subspace from which the initial ensemble was taken but their variance is scaled by a factor λ.

Figure 5.9 compares the performance of the bias-blind estimator for different magnitudes of process noise. The base case (without process noise) is represented by the green curve. One can marginally improve this performance by adding the right amount of process noise (red curve). This will increase the perceived reliability of the measurements, increasing the responsiveness of the estimator and reducing the observation data mismatch. Note, however, that the average error, both close to the body (solid lines) and away from it (dashed lines) does not change significantly. Also, this approach has its limits: too much noise will eventually dominate the estimator dynamics (blue curve). That notwithstanding, the bias-blind estimator final error is two orders of magnitude larger than the one that could be achieved in

(a) State error evolution.

(b) Observation error evolution.
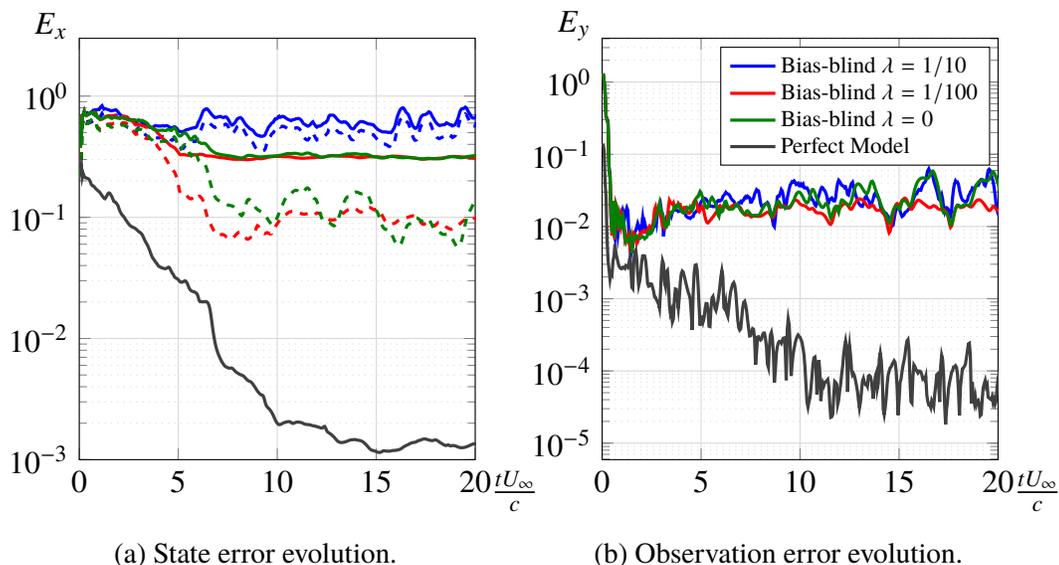
Figure 5.9: Bias-blind estimator performance highlighting the deleterious effect of the dynamics and observation bias ($R = 10^{-4}$). Black lines correspond to the mean square error evaluated in the entire computational domain, while the gray lines restrict this evaluation to the region outside an unit circle centered at the plate.

a perfect-model framework.

5.6 Bias-aware Estimation

In this section we evaluate the proposed bias-aware estimator performance when the exact flat plate bias statistics are used to form the POD-based Γ matrices. The AR1 model is chosen to represent the dynamics of the POD coefficients. Because the proposed scheme only adds $n_o + n_s = 35$ degrees of freedom to the much larger state vector x (about 15000 degrees of freedom), the additional computational cost is negligible. Figure 5.10a shows a 33% state error improvement for the entire domain, while the error far from the body improves 60%.

The bias dynamics are forced with process noise with covariance matrices $R^b = \lambda_o I_{n_o}$ and $Q^b = \lambda_s I_{n_s}$. The existence of process noise leads to a sustainably larger variance for the bias parameters, which allows for correction to be consistently made throughout the estimation window. This feature is especially important for problems like the present one, in which the bias cannot be considered slowly-varying. In fact, the bias is expected to exhibit a periodic behavior as the flow itself is periodic with the vortex-shedding period being the fundamental time scale. Figure 5.11 shows

(a) State error evolution.

(b) Observation error evolution.
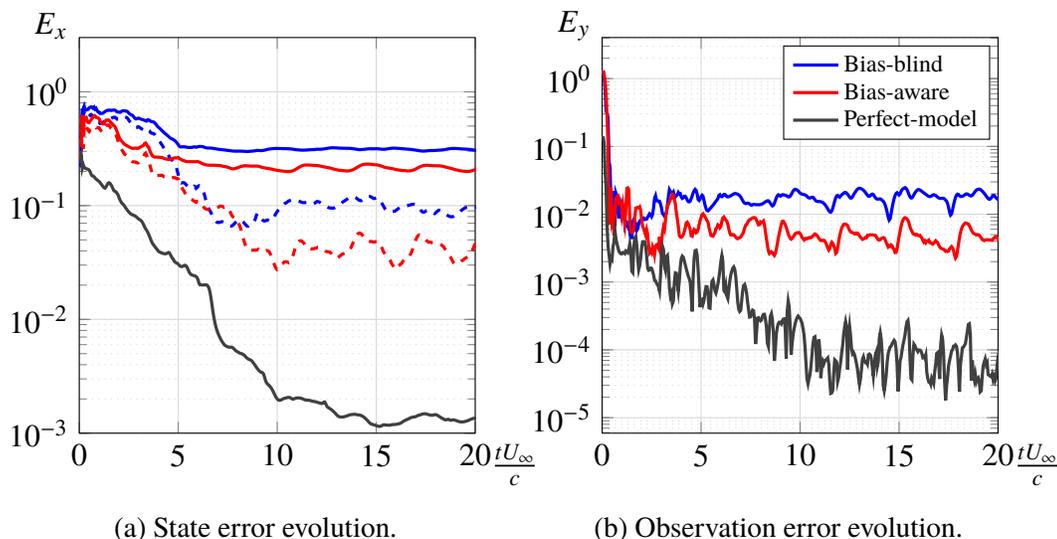
Figure 5.10: Bias-aware estimator performance ($R = 10^{-4}$, $\lambda_o = \sqrt{10}/10$ and $\lambda_s = \sqrt{10}/10$). Black lines correspond to the mean square error evaluated in the entire computational domain, while the gray lines restrict this evaluation to the region outside a unit circle centered at the plate.

how different choices of the noise magnitudes impact the state and observation error estimates. Larges values for the noise parameters favor smaller measurement mismatches (by allowing more aggressive analysis) at the expense of a possibly larger state error.

As Fig. 5.10b indicates, bias correction decreases the pressure error by 80%. Figure 5.12 displays an example of the correction introduced by the proposed scheme to the estimated output. Correction seems to be less effective near the leading edge, possibly because of the large pressure gradients that appear in there regions. As a consequence, global quantities like the lift coefficient also have their estimates improved.

5.7 Imperfect Bias Statistics

The biggest weakness of the proposed methodology is the need of a priori knowledge of the exact bias statistics, something that is rarely available. In order to show that good performance can also be achieved with imperfect statistics, an intermediary resolution ($Re_\Delta = 2$) is used in lieu of the fine-grid solution for the evaluation of the bias statistics. The performance of the resulting estimator (shown in Fig. 5.13) is very similar to the one obtained with the exact statistics. This seems to indicate

(a) State error evolution.

(b) Observation error evolution.
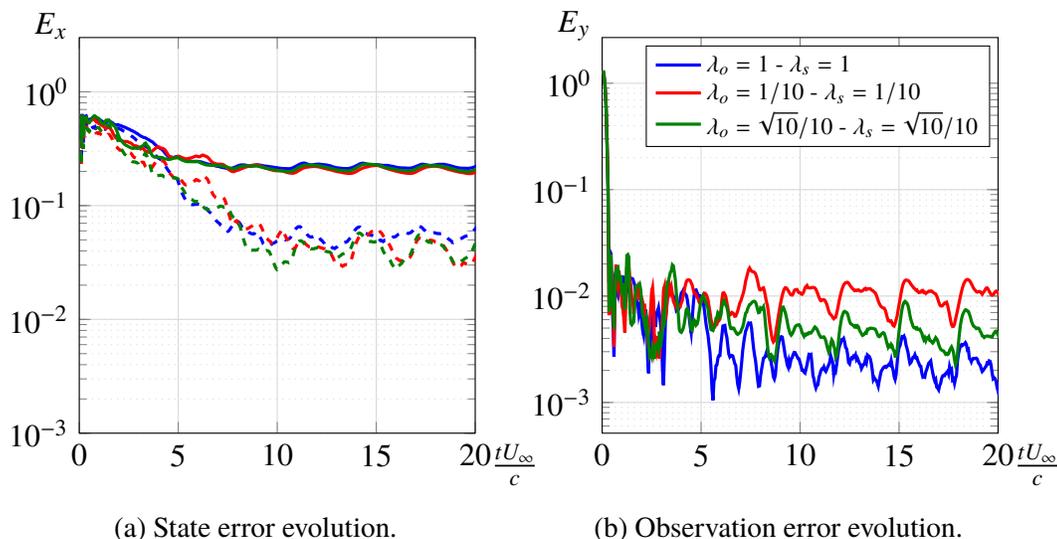
Figure 5.11: Effect of the magnitude of the process noise to bias dynamics ($R = 10^{-4}$). Solid lines correspond to the mean square error evaluated in the entire computational domain, while dashed lines restrict this evaluation to the region outside an unit circle centered at the plate.

that as long as one can estimate the structure of the bias, explicitly tracking it is beneficial for the estimation.

5.8 AR-based vs DMD-based Bias Models

We now address the effect of different choices of models for the bias dynamics on the performance of the bias-aware estimator. Figure 5.14 compares the POD-AR1 estimator presented in the previous sections with the DMD-based estimator. Recalling Fig. 5.7, we use 12 DMD modes to represent the observation bias and 18 DMD modes to represent the state forecast bias, which makes the cost comparable to the AR1 model setup. In terms of state error, both estimators have comparable performance, but the DMD-based estimator delivers poorer estimates for the estimated measurements.

5.9 Airfoil Case

Next we present the results of applying the bias-aware methodology to the airfoil case. As shown in section 5.4 (Figures 5.3 and 5.2), bias is more pronounced in the airfoil case. The bias-blind estimator displays a poorer performance in the airfoil case than in the flat plate case. The error is particularly high at early times in the
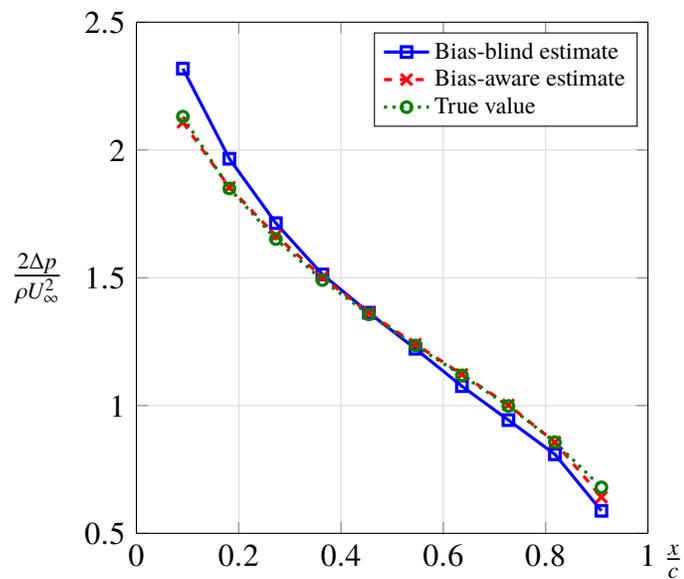
(a) Normal stresses.
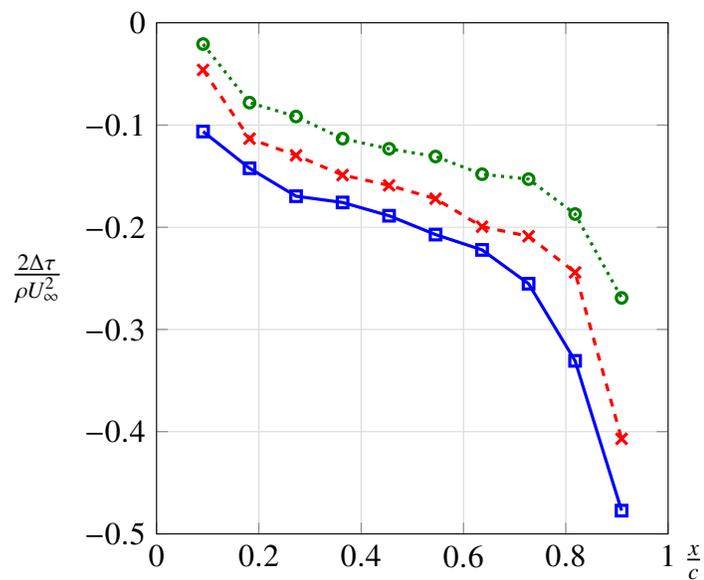
(b) Tangential stresses.

Figure 5.12: Estimated pressure on the surface of the flat plate at end of a simulation window ($tU_\infty/c = 20$).

(a) State error evolution.

(b) Observation error evolution.
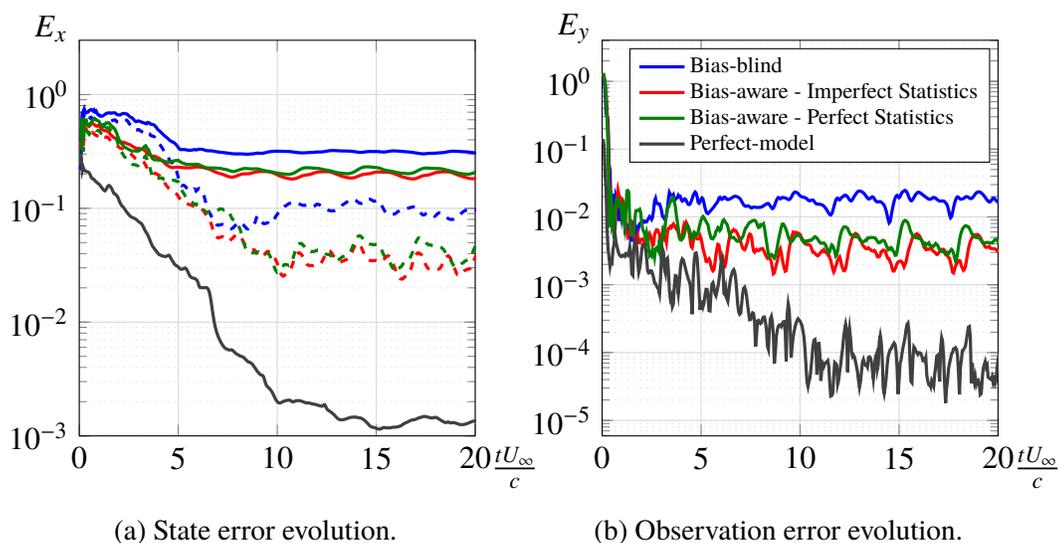
Figure 5.13: Bias-aware estimator performance with imperfect statistics ($R = 10^{-4}$). Solid lines correspond to the mean square error evaluated in the entire computational domain, while the dashed lines restrict this evaluation to the region outside a unit circle centered at the plate.

(a) State error evolution.

(b) Observation error evolution.
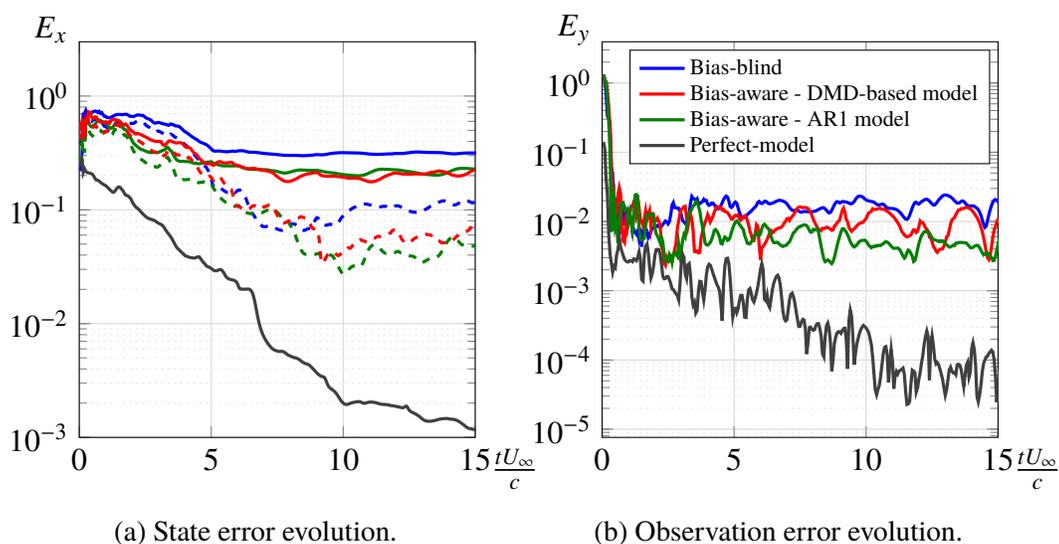
Figure 5.14: Effect of different choices of models for the bias dynamics on the performance of the bias-aware estimator. Solid lines correspond to the mean square error evaluated in the entire computational domain, while the dashed lines restrict this evaluation to the region outside a unit circle centered at the plate.

(a) State error evolution.

(b) Observation error evolution.
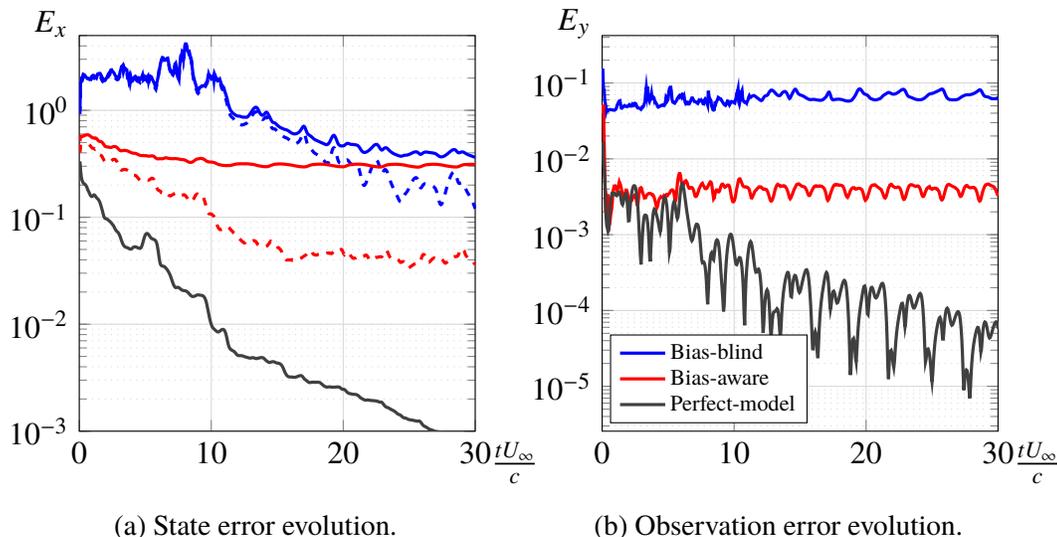
Figure 5.15: Bias-aware estimator performance ($R = 10^{-4}$, $\lambda_o = 1/10$ and $\lambda_s = 1$) for the inclined NACA 0009 problem. Solid lines correspond to the mean square error evaluated in the entire computational domain, while the dashed lines restrict this evaluation to the region outside an unit circle centered at the airfoil.

estimation window, when the measurement bias severely debilitates the accuracy of the corrections.

Again, we model the bias dynamics using the POD-AR1 model (with the same number of modes as before). Figure 5.15 shows that the bias-aware estimator greatly improves the accuracy throughout the entire estimation window, achieving a long-term error reduction of 85% for the state and 90% for the measurements. Figure 5.16 compares the estimated measurements to their real values before assimilation. The bias scheme is able to successfully correct the stresses on the surface of the airfoil. It can be noted, however, that most of the persistent error is located near the trailing edge.

5.10   Summary

In this chapter we addressed the problem of using an ensemble of under-resolved simulations to track the trajectory of the fully-resolved dynamical system. After showing that a bias-blind estimator could only go so far, we proposed an alternative method that improves the estimator performance by promoting bias awareness. Combining a low-order representation of the modeling error with a suitable bias dynamics, the methodology was able to improve the estimation accuracy of both

(a) Normal stresses.

(b) Tangential stresses.
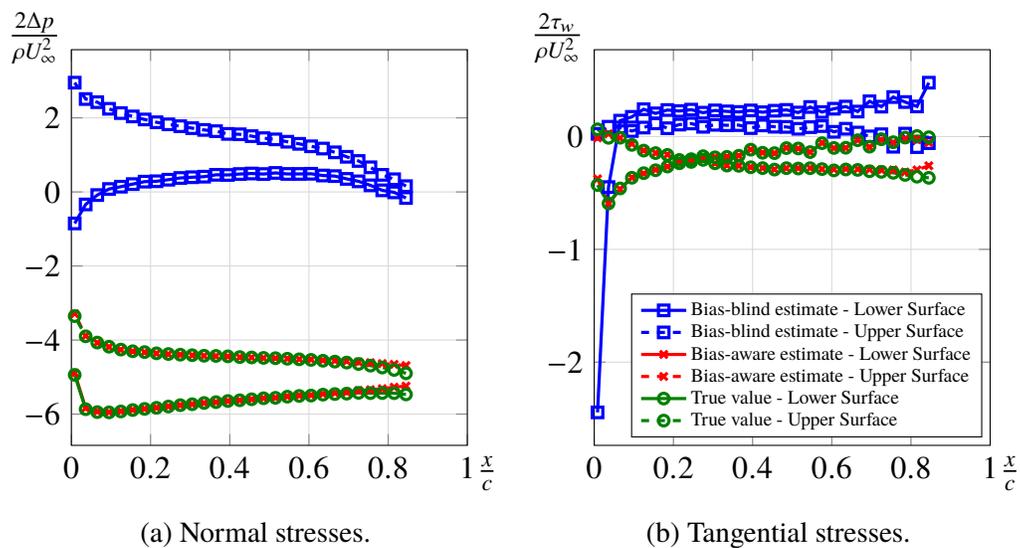
Figure 5.16: Estimated pressure on the surface of the NACA 0009 at the end of a simulation window ($tU_\infty/c = 30$).

the state and the output. Finally, we showed that the method's dependency on foreknowledge of the bias statistics can be alleviated by using extrapolated statistics.

Chapter 6

CONCLUSIONS AND FUTURE WORK

The main objective of this thesis was to obtain a better understanding of how ensemble methods can be used to address high-dimensional estimation problems in aerodynamics. More specifically we were interested in developing tools that would enable us to robustly assimilate data into a full-fledged CFD code.

A dual-experiment methodology, in which the numerical algorithm is used as both the estimator and as a surrogate for the measurements, was employed to investigate the performance of an EnKF-based estimator for aerodynamic problems. The estimator was demonstrated to provide a framework that combines the fidelity and robustness of a high-dimensional representation of the dynamical system with the computational efficiency of a low-order approximation of the state covariance.

While pursuing this goal, we were able to investigate several interesting aspects of the estimation framework and propose novel strategies for sensor placement and bias control. The EnKF-based estimator was successfully applied to the estimation of the flowfield and the aerodynamic force acting on bodies immersed in a low-Re flow, using measurements that reproduce the ones that are available in a laboratory setting, namely, pressure on the surfaces and velocity in the wake.

6.1 Perfect model framework

A reduced number of either velocity or pressure measurements was successfully used to estimate the phase of the vortex shedding, and thus the corresponding forces, in flow past a cylinder or an airfoil. In terms of performance, global measurements, such as the velocity in the wake, are more efficient in constraining the flow state. Estimation based on local measurements, such as pressure on the surface, results in longer transients before achieving the same state error level. However, using pressure measurements makes the estimator more responsive in tracking the forces acting on the body, which is especially desirable when freestream perturbations are present.

For flows like the ones considered, whose the dynamics evolve on a relatively small manifold, the results indicate that correspondingly small ensemble sizes are

sufficient to provide a suitable representation of the covariance matrix and lead to an asymptotic behavior of the estimator dynamics.

An analysis of the measurement's influence fields (the so-called representers) provided interesting *a posteriori* information about those measurements that were most effective during the filtering process. Thus, the representers provide useful guidance on sensor placement. In the cylinder case, the region of highest influence was coincident with the wavemaker region as predicted by structural sensitivity analysis. In the airfoil case, pressure measurements taken close to leading and trailing edges are most effective in the estimation process.

The effect of having a finite ensemble on the performance of the estimator, accelerating the collapse of the system covariance and generating spurious cross-correlations between the state variables, was discussed. In this work, the chosen form to mitigate this issue is by artificially inflating the covariance of the ensemble, a methodology usually referred to as covariance inflation. Later on in the thesis, covariance inflation was also used to weigh more the measurements in detriment to the forecast state as a first form of taking forecast errors into account (agnostic approach).

Because the pressure is a nonlinear function of the vorticity field, the effects of nonlinearities on the analysis step was discussed. Results indicated that, at least for the problem in question, Evensen's implicit linearization scheme outperforms iterative schemes that approximate the posterior conditional mean by the conditional mode.

The optimality of the estimation, as perceived from the time-correlation of the innovation sequence was evaluated. As expected, the estimator was able to optimally assimilate velocity measurements, which are a linear function of the state of the system, but not the pressure measurements, a nonlinear function of the state. Different schemes to deal with nonlinear observation functions were discussed and evaluated, and, surprisingly, the implicit linearization scheme first proposed by Evensen[19] demonstrated better performance than the iterative methods for the analyzed problems.

6.2   Parameter estimation

Assuming that the source of the forecast errors is known, parametric modeling can be used to upgrade the estimator model. The parameters and state variables can be jointly estimated by using a augmented state-vector approach. Additive

process noise can be introduced to ensure the responsiveness of the estimator to new disturbances. Using this approach, both the flow state and the lift were successfully estimated for an airfoil exposed to an unknown freestream velocity perturbation. Under a deterministic deceleration, the estimator accurately captured a steep change in the free-stream velocity (25% reduction of the initial velocity over 1 convective time unit). When subjected to a sequence of random gusts (reduced frequency up to 2 and amplitude up to 10% of the reference velocity), the estimator was able to achieve synchronization after a transient period of about 10 convective time units. Finally, the estimator was able to estimate the flow Reynolds number from pressure measurements taken on the surface of a flat plate, demonstrating that the estimator can be made robust to a range of initial conditions and uncertain parameters that would be problematic to standard reduced-order models typically used for flow control purposes.

6.3   A Bias-aware estimator

We also introduced a novel approach to use an EnKF framework to simultaneously mitigate the effects of biased forecast and observation models. In lieu of treating the bias as a single random variable, we split it into its slow and fast components. Colored noise processes were used to model the time correlation exhibited by the slow part of the bias, while the fast component is modeled by a white-noise process. The restriction of the bias dynamics to the low-rank subspace that contains most of the variance allows for a more efficient sampling of the state space and enables the use of fewer ensemble members to satisfactorily represent the system statistics.

The performance of the proposed estimator is assessed by employing an ensemble of coarse-grid simulations to track a fine-grid simulation of the low-Re flow past a flat plate at high angle of attack. Measurement data consists of pressure at ten different locations on the surface of the flat plate. The bias-aware estimator obtained about 70% state and observation error reduction for a marginal cost increment. Similar performance was achieved even when imperfect bias statistics where used.

6.4   Summary of Contributions

The main contributions presented in this thesis can be summarize as follows:

- Extension of the IBLGF framework to evaluate the pressure field from the vorticity field. Derivation of an IBLGF-based instantaneous mapping between

the vorticity field and the immersed boundary forces (Chapter 2).

- Survey of the EnKF capabilities under ideal conditions (perfect-model framework) which provided us a better understanding on how different parameters (covariance inflation, ensemble initialization, measurement noise levels and nonlinear observation schemes) affect the filter performance (Chapter 3).

- Optimal sensor placement based on the information contained in the representers matrix (Chapter 3).

- Demonstration of the EnKF capability of estimating the flow Reynold number and tracking rapidly changing freestream perturbation based on pressure measurements on the surface of the body (Chapter 4).

- Development of a bias-aware estimator based on a low-rank bias model which enables the user to use an ensemble of under-resolved simulations to track a fully-resolved reference data (Chapter 5).

Another contribution of this project which was not directly reported here is a result of cooperative effort with the UCLA and IIT. In Darakananda et al. [33], we combined the EnKF with a low-order vortex model in order to estimate the separated flow over a flat plate from pressure measurements on the surface of the plate. The estimator was demonstrated to be able not only of synchronizing with the reference data but also provide a good estimate of the aerodynamic force in the presence of sporadic incident gusts.

6.5   Recommendations for future work

As documented in this thesis, ensemble methods are a very efficient way of integrating complex models into estimators. An obvious path forward is to use the EnKF in conjunction with more complex flow configurations, which can include 2D cases with higher Reynolds numbers. Despite no longer being an accurate representation of reality, the resulting dynamics can display a chaotic behavior that may constitute a good challenge for the estimator.

A next step would then be to assimilate actual experimental data. Suitable models in this case would likely be three-dimensional and possibly require some appropriate turbulence treatment. The increased cost due to the larger number of degrees of

freedom will require extensive and efficient use of parallel methodologies in order to make real-time estimation possible.

Without doubt, the greater disadvantage of the proposed bias-aware method is the requirements of some form of a priori representation of the bias statistics. Although the results using extrapolated statistics are encouraging to the extent that they signal that a series of a priori simulations may be able to provide enough information, a methodology that forgoes the need for a priori statistics would be very welcome. A very promising direction points towards multi-fidelity estimators, which intend to leverage information obtained from different models (each of them with their own strengths and fragilities) to optimize the ratio accuracy/cost. Still along these lines, Multilevel Monte Carlo methods[105, 106] have been showing interesting results, employing a sequence of models of increasing complexity.

Finally, recalling what was exposed in section 1.2, flow estimation can be an end in itself (enhanced experiments), or an intermediary step to closed-loop flow control. A robust estimation strategy can be the missing technological piece to the development of the new generation of controllers.

BIBLIOGRAPHY

[1]   S. L. Brunton and B. R. Noack. "Closed-Loop Turbulence Control: Progress and Challenges". In: *Applied Mechanics Reviews* 67.5 (2015), p. 050801. DOI: 10.1115/1.4031175.

[2]   S. Ahuja and C. W. Rowley. "Feedback Control of Unstable Steady States of Flow Past a Flat Plate Using Reduced-Order Estimators". In: *Journal of Fluid Mechanics* 645 (Feb. 2010), p. 447. DOI: 10.1017/S0022112009992655.

[3]   T. L. B. Flinois and A. S. Morgans. "Feedback Control of Unstable Flows: a Direct Modelling Approach Using the Eigensystem Realisation Algorithm". In: *Journal of Fluid Mechanics* 793 (2016), p. 41. DOI: 10.1017/jfm.2016.111.

[4]   R. E. Kalman. "A New Approach to Linear Filtering and Prediction Problems". In: *Journal of basic Engineering* 82.1 (1960), pp. 35–45. DOI: 10.1115/1.3662552.

[5]   T. Hayase, K. Nisugi, and A. Shirai. "Numerical Realization for Analysis of Real Flows by Integrating Computation and Measurement". In: *International Journal for Numerical Methods in Fluids* 47.6-7 (2005), pp. 543–559. DOI: 10.1002/fld.829.

[6]   K. Nisugi, T. Hayase, and A. Shirai. "Fundamental Study of Hybrid Wind Tunnel Integrating Numerical Simulation and Experiment in Analysis of Flow Field". In: *JSME International Journal Series B* 47.3 (2004), pp. 593–604. DOI: 10.1299/jsmeb.47.593.

[7]   T. Hayase. "Numerical Simulation of Real-world Flows". In: *Fluid Dynamics Research* 47.5 (2015), p. 051201. DOI: 10.1088/0169-5983/47/5/051201.

[8]   J. M. Lewis et al. *Dynamic Data Assimilation: a Least Squares Approach*. Vol. 13. Cambridge University Press, 2006.

[9]   N. Papadakis and E. Memin. "Variational Assimilation of Fluid Motion from Image Sequence". In: *SIAM Journal on Imaging Sciences* 1.4 (2008), pp. 343–363. DOI: 10.1137/080713896.

[10] A. Gronskis, D. Heitz, and E. Memin. "Inflow and Initial Conditions for Direct Numerical Simulation Based on Adjoint Data Assimilation". In: *Journal of Computational Physics* 242 (2013), pp. 480–497. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2013.01.051.

[11] T. Suzuki, H. Ji, and F. Yamamoto. "Unsteady PTV Velocity Field Past an Airfoil Solved With DNS: Part 1. Algorithm of Hybrid Simulation and Hybrid Velocity Field at Re $\approx 10^3$". In: *Experiments in Fluids* 47.6 (June 2009), pp. 957–976. DOI: 10.1007/s00348-009-0691-9.

[12] T. Suzuki et al. "Unsteady PTV Velocity Field Past an Airfoil Solved With DNS: Part 2. Validation and Application at Reynolds Numbers Up to Re $\leq 10^4$". In: *Experiments in Fluids* 47.6 (June 2009), pp. 977–994. DOI: 10.1007/s00348-009-0692-8.

[13] D. P. G. Foures et al. "A Data-assimilation Method for Reynolds-averaged Navier-Stokes-driven Mean-Flow Reconstruction". In: *Journal of Fluid Mechanics* 759 (2014), pp. 404–431. DOI: 10.1017/jfm.2014.566.

[14] S. Symon et al. "Data Assimilation of Mean Velocity from 2D PIV Measurements of Flow Over an Idealized Airfoil". In: *Experiments in Fluids* 58.5 (Apr. 2017), p. 61. ISSN: 1432-1114. DOI: 10.1007/s00348-017-2336-8. URL: https://doi.org/10.1007/s00348-017-2336-8.

[15] J. Gerhard et al. "Model-Based Control of Vortex Shedding Using Low-Dimensional Galerkin Models". In: *Proceedings of the 33rd AIAA Fluid Dynamics Conference and Exhibit* (Orlando, FL). AIAA 2010-4262. AIAA. June 2003. DOI: 10.2514/6.2003-4262.

[16] K. Aleksic et al. "Robust Nonlinear Control Versus Linear Model Predictive Control of a Bluff Body Wake". In: *Proceedings of the 5th AIAA Flow Control Conference* (Chicago, IL). AIAA 2010-4833. AIAA. June 2010. DOI: 10.2514/6.2010-4833.

[17] I. Fukumori and P. Malanotte-Rizzoli. "An Approximate Kalman Filter for Ocean Data Assimilation: an Example With an Idealized Gulf Stream Model". In: *Journal of Geophysical Research: Oceans* 100.C4 (1995), pp. 6777–6793. DOI: 10.1029/94JC03084.

[18] T. Suzuki. "Reduced-Order Kalman-Filtered Hybrid Simulation Combining Particle Tracking Velocimetry and Direct Numerical Simulation". In: *Journal of Fluid Mechanics* 709 (2012), pp. 249–288. DOI: 10.1017/jfm.2012.334.

[19] G. Evensen. *Data Assimilation: the Ensemble Kalman Filter*. Springer Science & Business Media, 2009. ISBN: 3642037100.

[20] E. Kalnay. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 2003.

[21] J. Hoepffner et al. "State Estimation in Wall-bounded Flow Systems. Part 1. Perturbed Laminar Flows". In: *Journal of Fluid Mechanics* 534 (June 2005), pp. 263–294. DOI: 10.1017/S0022112005004210.

[22] M. Chevalier et al. "State Estimation in Wall-Bounded Flow Systems. Part 2. Turbulent Flows". In: *Journal of Fluid Mechanics* 552.-1 (Mar. 2006), p. 167. DOI: 10.1017/S0022112005008578.

[23] C. H. Colburn, J. B. Cessna, and T. R. Bewley. "State Estimation in Wall-Bounded Flow Systems. Part 3. The Ensemble Kalman Filter". In: *Journal of Fluid Mechanics* 682 (Aug. 2011), pp. 289–303. DOI: `10.1017/jfm.2011.222`.

[24] H. Kato and S. Obayashi. "Hybrid Wind Tunnel Based on Ensemble Kalman Filter". In: *14th International Conference on Information Fusion*. IEEE, July 2011, pp. 1–8.

[25] R. Kikuchi, T. Misaka, and S. Obayashi. "Assessment of Probability Density Function Based on POD Reduced-Order Model for Ensemble-Based Data Assimilation". In: *Fluid Dynamics Research* 47.5 (2015), p. 051403. DOI: `10.1088/0169-5983/47/5/051403`.

[26] H. Kato et al. "A Data Assimilation Methodology for Reconstructing Turbulent Flows Around Aircraft". In: *Journal of Computational Physics* 283.C (2015), pp. 559–581. ISSN: 0021-9991. DOI: `10.1016/j.jcp.2014.12.013`.

[27] V. Mons et al. "Reconstruction of Unsteady Viscous Flows Using Data Assimilation Schemes". In: *Journal of Computational Physics* 316 (2016), pp. 255–280. ISSN: 0021-9991. DOI: `10.1016/j.jcp.2016.04.022`.

[28] T. R. Bewley and B. Protas. "Skin Friction and Pressure: the Footprints of Turbulence". In: *Physica D: Nonlinear Phenomena* 196.1 (2004), pp. 28–44. ISSN: 0167-2789. DOI: `10.1016/j.physd.2004.02.008`.

[29] P. Ruhnau, A. Stahl, and C. Schnorr. "Variational Estimation of Experimental Fluid Flows With Physics-based Spatio-temporal Regularization". In: *Measurement Science and Technology* 18.3 (2007), p. 755. DOI: `10.1088/0957-0233/18/3/027`.

[30] V. Mons, J.-C. Chassaing, and P. Sagaut. "Optimal Sensor Placement for Variational Data Assimilation of Unsteady Flows Past a Rotationally Oscillating Cylinder". In: *Journal of Fluid Mechanics* 823 (2017), pp. 230–277. DOI: `10.1017/jfm.2017.313`.

[31] M. Meldi and A. Poux. "A Reduced Order Model Based on Kalman Filtering for Sequential Data Assimilation of Turbulent Flows". In: *Journal of Computational Physics* 347 (2017), pp. 207–234. ISSN: 0021-9991. DOI: `10.1016/j.jcp.2017.06.042`.

[32] S. P. Symon. "Reconstruction and Estimation of Flows Using Resolvent Analysis and Data-Assimilation". PhD Thesis. California Institute of Technology, May 11, 2018. DOI: `10.7907/B4K7-K876`. URL: `http://resolver.caltech.edu/CaltechTHESIS:05302018-181049042` (visited on 09/06/2018).

[33] D. Darakananda et al. "EnKF-based Dynamic Estimation of Separated Flows with a Low-Order Vortex Model". In: *2018 AIAA Aerospace Sciences Meeting*. AIAA 2018-0811. Reston, Virginia: American Institute of Aeronautics and Astronautics, Jan. 2018. DOI: `10.2514/6.2018-0811`.

[34] A. F. C. da Silva and T. Colonius. "Ensemble-Based State Estimator for Aerodynamic Flows". In: *AIAA Journal* (June 2018), pp. 1–11. DOI: `10.2514/1.J056743`.

[35] A. F. C. da Silva and T. Colonius. "A Bias-aware EnKF Estimator for Aerodynamic Flows". In: *48th AIAA Fluid Dynamics Conference*. AIAA 2018-3225. Reston, Virginia: American Institute of Aeronautics and Astronautics, June 2018. DOI: `10.2514/6.2018-3225`.

[36] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Courier Corporation, 1970.

[37] A. S. Charles. *Kalman Filtering: a Bayesian Approach*. Dec. 14, 2017. URL: `http://adamsc.mycpanel.princeton.edu/documents/KalmanFilterBayesDerivation.pdf` (visited on 08/01/2018).

[38] P. J. van Leeuwen. "Particle Filtering in Geophysical Systems". In: *Monthly Weather Review* 137.12 (2009), pp. 4089–4114. DOI: `10.1175/2009MWR2835.1`.

[39] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Science & Business Media, 2013.

[40] P. L. Houtekamer and F. Zhang. "Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation". In: *Monthly Weather Review* 144.12 (2016), pp. 4489–4532. DOI: `10.1175/MWR-D-15-0440.1`.

[41] R. F. Stengel. *Optimal Control and Estimation*. Courier Corporation, 2012.

[42] B. D. O. Anderson and J. B. Moore. *Optimal Filtering*. Courier Corporation, 2012.

[43] K. Law, A. Stuart, and K. Zygalakis. *Data Assimilation: A Mathemathical Introduction*. Cham: Springer, 2015. ISBN: 978-3-319-20324-9. DOI: `10.1007/978-3-319-20325-6`.

[44] S. Gillijns et al. "What is the Ensemble Kalman Filter and How Well Does it Work?" In: *Proceedings of the 2006 American Control Conference*. IEEE. June 2006, pp. 4448–4453. DOI: `10.1109/ACC.2006.1657419`.

[45] A. Gelb. *Applied Optimal Estimation*. MIT press, 1974.

[46] S. J. Julier and J. K. Uhlmann. "Unscented Filtering and Nonlinear Estimation". In: *Proceedings of the IEEE* 92.3 (2004), pp. 401–422. DOI: `10.1109/JPROC.2003.823141`.

[47]   E. A. Wan and R. van Der Merwe. "The Unscented Kalman Filter for Non-linear Estimation". In: *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*. IEEE. 2000, pp. 153–158. DOI: `10.1109/ASSPCC.2000.882463`.

[48]   A. C. Lorenc. "Analysis Methods for Numerical Weather Prediction". In: *Quarterly Journal of the Royal Meteorological Society* 112.474 (1986), pp. 1177–1194. ISSN: 1477-870X. DOI: `10.1002/qj.49711247414`. URL: `10.1002/qj.49711247414`.

[49]   R. E. Bellman. *Dynamic Programming*. Dover Publications, Incorporated, 2003. ISBN: 0486428095.

[50]   P. J. van Leeuwen. "Particle Filtering in Geophysical Systems". In: *Monthly Weather Review* 137.12 (2009), pp. 4089–4114. DOI: `10.1175/2009MWR 2835.1`.

[51]   G. Evensen. "Sequential Data Assimilation With a Nonlinear Quasi-Geostrophic Model Using Monte Carlo Methods to Forecast Error Statistics". In: *Journal of Geophysical Research: Oceans* 99.C5 (1994), pp. 10143–10162. DOI: `10.1029/94JC0057`.

[52]   T. Bengtsson, C. Snyder, and D. Nychka. "Toward a Nonlinear Ensemble Filter for High-Dimensional Systems". In: *Journal of Geophysical Research: Atmospheres* 108.D24 (2003). DOI: `10.1029/2002JD002900`.

[53]   G. Evensen. "Sampling Strategies and Square Root Analysis Schemes for the EnKF". In: *Ocean Dynamics* 54.6 (2004), pp. 539–560. DOI: `10.1007/s10236-004-0099-2`.

[54]   J. L. Anderson and S. L. Anderson. "A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts". In: *Monthly Weather Review* 127.12 (1999), pp. 2741–2758. DOI: `10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2`.

[55]   N. Papadakis et al. "Data Assimilation With the Weighted Ensemble Kalman Filter". In: *Tellus A* 62.5 (2010), pp. 673–697. DOI: `10.1111/j.1600-0870.2010.00461.x`.

[56]   P. Sakov, G. Evensen, and L. Bertino. "Asynchronous Data Assimilation with the EnKF". In: *Tellus A* 62.1 (2009), pp. 24–29. DOI: `10.1111/j.1600-0870.2009.00417.x`.

[57]   G. Evensen and P. J. van Leeuwen. "Assimilation of Geosat Altimeter Data for the Agulhas Current Using the Ensemble Kalman Filter With a Quasi-geostrophic Model". In: *Monthly Weather Review* 124.1 (1996), pp. 85–96. DOI: `10.1175/1520-0493(1996)124<0085:AOGADF>2.0.CO;2`.

[58]   W. G. Lawson and J. A. Hansen. "Implications of Stochastic and Deterministic Filters as Ensemble-Based Data Assimilation Methods in Varying Regimes of Error Growth". In: *Monthly Weather Review* 132.8 (2004), pp. 1966–1981. DOI: `10.1175/1520-0493(2004)132<1966:IOSADF>2.0.CO;2`.

[59]   F. Le Gland, V. Monbet, and V. Tran. *Large Sample Asymptotics for the Ensemble Kalman Filter*. Research Report RR-7014. Institut National de Recherche en Informatique et en Automatique (INRIA), 2009, p. 25. URL: `https://hal.inria.fr/inria-00409060`.

[60]   G. Evensen. "The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation". In: *Ocean dynamics* 53.4 (2003), pp. 343–367. DOI: `10.1007/s10236-003-0036-9`.

[61]   M. Bocquet. "Ensemble Kalman Filtering Without the Intrinsic Need for Inflation". In: *Nonlinear Processes in Geophysics* 18.5 (2011), pp. 735–750. DOI: `10.5194/npg-18-735-2011`. URL: `https://www.nonlin-processes-geophys.net/18/735/2011/`.

[62]   P. L. Houtekamer and H. L. Mitchell. "Data Assimilation Using an Ensemble Kalman Filter Technique". In: *Monthly Weather Review* 126.3 (1998), pp. 796–811. DOI: `10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2`.

[63]   P. J. van Leeuwen. "Comment on "Data Assimilation Using an Ensemble Kalman Filter Technique"". In: *Monthly Weather Review* 127.6 (1999), pp. 1374–1377. DOI: `10.1175/1520-0493(1999)127<1374:CODAUA>2.0.CO;2`.

[64]   M. Bocquet and P. Sakov. "Combining Inflation-free and Iterative Ensemble Kalman Filters for Strongly Nonlinear Systems". In: *Nonlinear Processes in Geophysics* 19.3 (2012), pp. 383–399. DOI: `10.5194/npg-19-383-2012`. URL: `https://www.nonlin-processes-geophys.net/19/383/2012/`.

[65]   M. Bocquet, P. N. Raanes, and A. Hannart. "Expanding the validity of the Ensemble Kalman Filter Without the Intrinsic Need for Inflation". In: *Nonlinear Processes in Geophysics* 22.6 (2015), pp. 645–662. DOI: `10.5194/npg-22-645-2015`. URL: `https://www.nonlin-processes-geophys.net/22/645/2015/`.

[66]   T. M. Hamill, J. S. Whitaker, and C. Snyder. "Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter". In: *Monthly Weather Review* 129.11 (2001), pp. 2776–2790. DOI: `10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2`.

[67]   D. T. B. Kelly, K. J. H. Law, and A. M. Stuart. "Well-Posedness and Accuracy of the Ensemble Kalman Filter in Discrete and Continuous Time". In: *Nonlinearity* 27.10 (2014), p. 2579. DOI: `10.1088/0951-7715/27/10/2579`.

[68]   J. S. Whitaker and T. M. Hamill. "Evaluating Methods to Account for System Errors in Ensemble Data Assimilation". In: *Monthly Weather Review* 140.9 (2012), pp. 3078–3089. DOI: `10.1175/MWR-D-11-00276.1`.

[69]   X. Luo and I. Hoteit. "Robust Ensemble Filtering and its Relation to Covariance Inflation in the Ensemble Kalman Filter". In: *Monthly Weather Review* 139.12 (2011), pp. 3938–3953. DOI: `10.1175/MWR-D-10-05068.1`. eprint: `https://doi.org/10.1175/MWR-D-10-05068.1`. URL: `https://doi.org/10.1175/MWR-D-10-05068.1`.

[70]   W. Sacher and P. Bartello. "Sampling Errors in Ensemble Kalman Filtering. Part I: Theory". In: *Monthly Weather Review* 136.8 (2008), pp. 3035–3049. DOI: `10.1175/2007MWR2323.1`.

[71]   B. M. Bell and F. W. Cathey. "The iterated Kalman filter update as a Gauss-Newton method". In: *IEEE Transactions on Automatic Control* 38.2 (Feb. 1993), pp. 294–297. ISSN: 0018-9286. DOI: `10.1109/9.250476`.

[72]   A. Bjorck. "9. Nonlinear Least Squares Problems". In: *Numerical Methods for Least Squares Problems*. 1996. Chap. 9, pp. 339–358. DOI: `10.1137/1.9781611971484.ch9`. eprint: `https://epubs.siam.org/doi/pdf/10.1137/1.9781611971484.ch9`.

[73]   B. M. Bell and F. W. Cathey. "The iterated Kalman filter update as a Gauss-Newton method". In: *IEEE Transactions on Automatic Control* 38.2 (Feb. 1993), pp. 294–297. ISSN: 0018-9286. DOI: `10.1109/9.250476`.

[74]   S. E. Cohn. "An Introduction to Estimation Theory". In: *Journal of the Meteorological Society of Japan. Ser. II* 75.1B (1997), pp. 257–288. DOI: `10.2151/jmsj1965.75.1B_257`.

[75]   M. Zupanski. "Maximum Likelihood Ensemble Filter: Theoretical Aspects". In: *Monthly Weather Review* 133.6 (2005), pp. 1710–1726. DOI: `10.1175/MWR2946.1`.

[76]   C. H. Bishop, B. J. Etherton, and S. J. Majumdar. "Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects". In: *Monthly Weather Review* 129.3 (2001), pp. 420–436. DOI: `10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2`.

[77]   Y. Gu, D. S. Oliver, et al. "An iterative ensemble Kalman filter for multiphase fluid flow data assimilation". In: *SPE Journal* 12.04 (2007), pp. 438–446. DOI: `10.2118/108438-PA`.

[78]   K. Taira and T. Colonius. "The Immersed Boundary Method: a Projection Approach". In: *Journal of Computational Physics* 225.2 (2007), pp. 2118–2137. DOI: `10.1016/j.jcp.2007.03.005`.

[79] T. Colonius and K. Taira. "A Fast Immersed Boundary Method Using a Nullspace Approach and Multi-Domain Far-Field Boundary Conditions". In: *Computer Methods in Applied Mechanics and Engineering* 197.25 (2008), pp. 2131–2146. DOI: `10.1016/j.cma.2007.08.014`.

[80] S. Liska and T. Colonius. "A Parallel Fast Multipole Method for Elliptic Difference Equations". In: *Journal of Computational Physics* 278 (2014), pp. 76–91. DOI: `10.1016/j.jcp.2014.07.048`.

[81] S. Liska and T. Colonius. "A Fast Immersed Boundary Method for External Incompressible Viscous Flows Using Lattice Green's Functions". In: *Journal of Computational Physics* 331 (2017), pp. 257–279. DOI: `10.1016/j.jcp.2016.11.034`. URL: `http://colonius.caltech.edu/pdfs/LiskaColonius2017.pdf`.

[82] H.-C. Tsai and T. Colonius. "Coriolis Effect on Dynamic Stall in a Vertical Axis Wind Turbine". In: *AIAA Journal* 54.1 (Jan. 2016), pp. 216–226. DOI: `10.2514/1.J054199`.

[83] H. C. Tsai. "Numerical Investigation of Vertical-Axis Wind Turbines at Low Reynolds Number". Thesis (Ph.D.) California Institute of Technology, 2016. DOI: `10.7907/Z9SF2T5R`.

[84] X. Yang et al. "A Smoothing Technique for Discrete Delta Functions With Application to Immersed Boundary Method in Moving Boundary Simulations". In: *Journal of Computational Physics* 228.20 (2009), pp. 7821–7836. ISSN: 0021-9991. DOI: `10.1016/j.jcp.2009.07.023`.

[85] A. Goza et al. "Accurate Computation of Surface Stresses and Forces With Immersed Boundary Methods". In: *Journal of Computational Physics* 321 (2016), pp. 860–873. DOI: `10.1016/j.jcp.2016.06.014`.

[86] A. M. Roma, C. S. Peskin, and M. J. Berger. "An Adaptive Version of the Immersed Boundary Method". In: *Journal of Computational Physics* 153.2 (1999), pp. 509–534. ISSN: 0021-9991. DOI: `10.1006/jcph.1999.6293`.

[87] Y. He and J. Li. "Numerical Implementation of the Crank-Nicolson/Adams-Bashforth Scheme for the Time-Dependent Navier-Stokes Equations". In: *International Journal for Numerical Methods in Fluids* 62.6 (2009), pp. 647–659. DOI: `10.1002/fld.2035`.

[88] S. Liska and T. Colonius. "A Fast Lattice Green's Function Method for Solving Viscous Incompressible Flows on Unbounded Domains". In: *Journal of Computational Physics* 316 (2016), pp. 360–384. DOI: `10.1016/j.jcp.2016.04.023`. URL: `http://colonius.caltech.edu/pdfs/LiskaColonius2016.pdf`.

[89] U. Fey, M. Konig, and H. Eckelmann. "A New Strouhal-Reynolds-number Relationship for the Circular Cylinder in the Range $47 < Re < 2 \times 10^5$". In: *Physics of Fluids* 10.7 (1998), pp. 1547–1549. DOI: `10.1063/1.869675`.

[90] C. H. K. Williamson. "The Existence of two Stages in the Transition to Three-Dimensionality of a Cylinder Wake". In: *Physics of Fluids* 31.11 (1988), pp. 3165–3168. DOI: 10.1063/1.866925.

[91] D. Sanz-Alonso and A. M. Stuart. "Long-time Asymptotics of the Filtering Distribution for Partially Observed Chaotic Dynamical Systems". In: *SIAM/ASA Journal on Uncertainty Quantification* 3.1 (2015), pp. 1200–1220. DOI: 10.1137/140997336.

[92] F. Giannetti and P. Luchini. "Structural Sensitivity of the First Instability of the Cylinder Wake". In: *Journal of Fluid Mechanics* 581 (June 2007), pp. 167–197. DOI: 10.1017/S0022112007005654.

[93] A. F. C. da Silva and T. Colonius. "An EnKF-based Flow State Estimator for Aerodynamic Flows". In: *8th AIAA Theoretical Fluid Mechanics Conference*. Reston, Virginia: American Institute of Aeronautics and Astronautics, June 2017. DOI: 10.2514/6.2017-3483.

[94] P. Matisko and V. Havlena. "Optimality Tests and Adaptive Kalman Filter". In: *IFAC Proceedings Volumes* 45.16 (2012). 16th IFAC Symposium on System Identification, pp. 1523–1528. ISSN: 1474-6670. DOI: 10.3182/20120711-3-BE-2027.00011.

[95] T. M. Pukkila and P. R. Krishnaiah. "On the use of autoregressive order determination criteria in multivariate white noise tests". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 36.9 (Sept. 1988), pp. 1396–1403. DOI: 10.1109/29.90367.

[96] J. L. Anderson. "An Ensemble Adjustment Kalman Filter for Data Assimilation". In: *Monthly Weather Review* 129.12 (2001), pp. 2884–2903. DOI: 10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

[97] S. Brunton et al. "Compressive Sensing and Low-Rank Libraries for Classification of Bifurcation Regimes in Nonlinear Dynamical Systems". In: *SIAM Journal on Applied Dynamical Systems* 13.4 (2014), pp. 1716–1732. DOI: 10.1137/130949282.

[98] B. Friedland. "Treatment of Bias in Recursive Filtering". In: *IEEE Transactions on Automatic Control* 14.4 (Aug. 1969), pp. 359–367. ISSN: 0018-9286. DOI: 10.1109/TAC.1969.1099223.

[99] D. P. Dee and A. M. Da Silva. "Data Assimilation in the Presence of Forecast Bias". In: *Quarterly Journal of the Royal Meteorological Society* 124.545 (1998), pp. 269–295. ISSN: 1477-870X. DOI: 10.1002/qj.49712454512.

[100] J.-P. Drecourt, H. Madsen, and D. Rosbjerg. "Bias Aware Kalman Filters: Comparison and Improvements". In: *Advances in Water Resources* 29.5 (2006), pp. 707–718. ISSN: 0309-1708. DOI: 10.1016/j.advwatres.2005.07.006.

[101] D. P. Dee. "On-line Estimation of Error Covariance Parameters for Atmospheric Data Assimilation". In: *Monthly Weather Review* 123.4 (1995), pp. 1128–1145. DOI: `10.1175/1520-0493(1995)123<1128:OLEOEC>2.0.CO;2`.

[102] S. E. Cohn and D. F. Parrish. "The Behavior of Forecast Error Covariances for a Kalman Filter in two Dimensions". In: *Monthly Weather Review* 119.8 (1991), pp. 1757–1785. DOI: `10.1175/1520-0493(1991)119<1757:TBOFEC>2.0.CO;2`.

[103] C. K. Chui and G. Chen. *Kalman Filtering With Real Time Applications*. Springer, 2009. DOI: `10.1007/978-3-319-47612-4`.

[104] P. J. Schmid. "Dynamic Mode Decomposition of Numerical and Experimental Data". In: *Journal of fluid mechanics* 656 (2010), pp. 5–28.

[105] M. B. Giles. "Multilevel Monte Carlo Methods". In: *Acta Numerica* 24 (2015), pp. 259–328.

[106] H. Hoel, K. J. H. Law, and R. Tempone. "Multilevel Ensemble Kalman Filtering". In: *SIAM Journal on Numerical Analysis* 54.3 (2016), pp. 1813–1839. DOI: `10.1137/15M100955X`.

[107] O. Buneman. "Analytic Inversion of the Five-point Poisson Operator". In: *Journal of Computational Physics* 8.3 (1971), pp. 500–505. ISSN: 0021-9991. DOI: `10.1016/0021-9991(71)90029-5`.

[108] A. Gillman and P. G. Martinsson. "A Fast Solver for Poisson Problems on Infinite Regular Lattices". In: *Journal of Computational and Applied Mathematics* 258 (2014), pp. 42–56. ISSN: 0377-0427. DOI: `10.1016/j.cam.2013.09.003`.

A p p e n d i x   A

IMPLICIT LINEARIZATION

The scheme proposed by Evensen [19] to deal with nonlinear observation functions can be understood as an approximate linearization about the ensemble-averaged state. To see this, notice that each of the columns of $H\hat{A}_k$ is given by

$$
\sqrt{q-1}\left\{H\hat{A}_k^w\right\}_j = h\left(\hat{z}_k^{(j)}\right) - \frac{1}{q}\sum_{i=1}^{q} h\left(\hat{z}_k^{(i)}\right)
$$

$$
= h\left(\hat{z}_k^{(j)}\right) - \frac{1}{q}\sum_{i=1}^{q}\left[h(\bar{\hat{z}}_k) + \frac{\partial h}{\partial z}(\bar{\hat{z}}_k)\left(\hat{z}_k^{(j)} - \bar{\hat{z}}_k\right) + O[(\hat{z}_k - \bar{\hat{z}}_k)^2]\right]
$$

$$
= h\left(\hat{z}_k^{(j)}\right) - h\left(\bar{\hat{z}}_k\right) + O[(\hat{z}_k - \bar{\hat{z}}_k)^2]
$$

$$
= \frac{\partial h}{\partial z}(\bar{\hat{z}}_k)\left(\hat{z}_k^{(j)} - \bar{\hat{z}}_k\right) + O[(\hat{z}_k - \bar{\hat{z}}_k)^2], \tag{A.1a}
$$

and, therefore,

$$
\hat{C}_k^{yy} = \frac{1}{q-1}\sum_{j=1}^{q}\left(h(\hat{z}_k^{(j)}) - \frac{1}{q}\sum_{i=1}^{q}h(\hat{z}_k^{(i)})\right)\left(h(\hat{z}_k^{(j)}) - \frac{1}{q}\sum_{i=1}^{q}h(\hat{z}_k^{(i)})\right)^T
$$

$$
= \left[\frac{\partial h}{\partial z}(\bar{\hat{z}}_k)\right]\left(\frac{1}{q-1}\sum_{j=1}^{q}\left(\hat{z}_k^{(j)} - \bar{\hat{z}}_k\right)\left(\hat{z}_k^{(j)} - \bar{\hat{z}}_k\right)^T\right)\left[\frac{\partial h}{\partial z}(\bar{\hat{z}}_k)\right]^T + O[(\hat{z}_k - \bar{\hat{z}}_k)^3]
$$

$$
= \left[\frac{\partial h}{\partial z}(\bar{\hat{z}}_k)\right]\hat{C}_k^{zz}\left[\frac{\partial h}{\partial z}(\bar{\hat{z}}_k)\right]^T + O[(\hat{z}_k - \bar{\hat{z}}_k)^3] \tag{A.2}
$$

and

$$
\hat{C}_k^{zy} = \hat{C}_k^{zz}\left[\frac{\partial h}{\partial z}(\bar{\hat{z}}_k)\right]^T + O[(\hat{z}_k - \bar{\hat{z}}_k)^3]. \tag{A.3}
$$

These approximations are consistent with the approximate minimization of an effective cost function

$$
J(z) = \frac{1}{2\alpha}\|z - \hat{z}_k^{(j)}\|_{\hat{C}^{zz}}^2 + \frac{1}{2}\left\|y_k - h(\bar{\hat{z}}) - \left[\frac{\partial h}{\partial z}(\bar{\hat{z}})\right](z - \bar{\hat{z}}) - v_k^{(j)}\right\|_R^2, \tag{A.4}
$$

which corresponds to the cost function of an extended Kalman filter in which the linearization is performed about the ensemble average instead of the prior state of the particle.

If $h(x)$ is linear, i.e. $h(x) = Hx$, then

$$
\begin{aligned}
\hat{C}_k^{yy} &= \frac{1}{q-1} \sum_{j=1}^{q} \left(H\hat{x}_k^j - \frac{1}{q} \sum_{i=1}^{q} H\hat{x}_k^i \right) \left(H\hat{x}_k^j - \frac{1}{q} \sum_{i=1}^{q} H\hat{x}_k^i \right)^T \\
&= \frac{1}{q-1} \sum_{j=1}^{q} H \left(\hat{x}_k^j - \frac{1}{q} \sum_{i=1}^{q} \hat{x}_k^i \right) \left(\hat{x}_k^j - \frac{1}{q} \sum_{i=1}^{q} \hat{x}_k^i \right)^T H^T \\
&= H\hat{C}_k^{xx} H^T
\end{aligned}
$$

(A.5)

and

$$
\hat{C}_k^{xy} = \frac{1}{q-1} \sum_{j=1}^{q} \left(\hat{x}_k^j - \frac{1}{q} \sum_{i=1}^{q} \hat{x}_k^i \right) \left(H\hat{x}_k^j - \frac{1}{q} \sum_{i=1}^{q} H\hat{x}_k^i \right)^T = \hat{C}_k^{xx} H^T ,
$$

(A.6)

and, thus, the standard EnKF analysis formula is recovered.

A p p e n d i x B

LATTICE GREEN'S FUNCTION FOR THE 2D LAPLACE OPERATOR

Lattice Green's functions (LGF) can be used to efficiently solve Poisson problems defined on the infinite integer lattice $\mathbb{Z}^n$, i.e.,

$$Lu(\mathbf{m}) = f(\mathbf{m}) \qquad \mathbf{m} \in \mathbb{Z}^n \,, \tag{B.1}$$

where L is the second-order discrete Laplace operator in $\mathbb{Z}^n$. In two dimensions ($n = 2$, $\mathbf{m} = (m_1, m_2)$), this operator is defined as

$$Lu(m_1, m_2) = 4u(m_1, m_2)$$
$$- u(m_1 + 1, m_2) - u(m_1 - 1, m_2) - u(m_1, m_2 + 1) - u(m_1, m_2 - 1) \,. \tag{B.2}$$

The solution of this operator for a single unitary charge placed at the origin is the corresponding lattice Green's function $\phi(m_1, m_2)$. In terms of this fundamental solution, Eq. B.1 has an explicit analytic solution in the form

$$u(\mathbf{m}) = [\phi \star f](\mathbf{m}) = \sum_{\mathbf{p} \in \mathbb{Z}^2} \phi(\mathbf{m} - \mathbf{p}) f(\mathbf{p}) \,. \tag{B.3}$$

Analytically, the LGF can be evaluated via the normalized Fourier integral

$$\phi(m_1, m_2) = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{\cos(t_1 m_1 + t_2 m_2) - 1}{4 \sin^2(t_1/2) + 4 \sin^2(t_2/2)} dt_1 dt_2 \,. \tag{B.4}$$

Numerically, however, the alternative expression derived by Buneman [107] is more stable

$$\phi(m, n) = -\frac{1}{2\pi} \Re \int_0^1 \left[1 - \left(\frac{\tau - \sqrt{i}}{\tau + \sqrt{i}} \right)^{m_1 + |m_2|} \left(\frac{\tau + \sqrt{-i}}{\tau - \sqrt{-i}} \right)^{m_1 - |m_2|} \right] \frac{d\tau}{\tau} \,. \tag{B.5}$$

Away from the origin, the following asymptotic expansion can be used instead [108]

$$\phi(m_1, m_2) = -\frac{1}{2\pi} \left(\log |\mathbf{m}| + \gamma + \frac{\log 8}{2} \right) + \frac{1}{24\pi} \frac{m_1^4 - 6m_1 m_2 + m_2^4}{|\mathbf{m}|^6}$$
$$+ \frac{1}{480\pi} \frac{43m_1^8 - 772m_1^6 m_2 + 1570m_1^4 m_2^4 - 772m_1^2 m_2^6 + 43m_2^8}{|\mathbf{m}|^{12}} + O(|\mathbf{m}|^{-6}) \,, \tag{B.6}$$
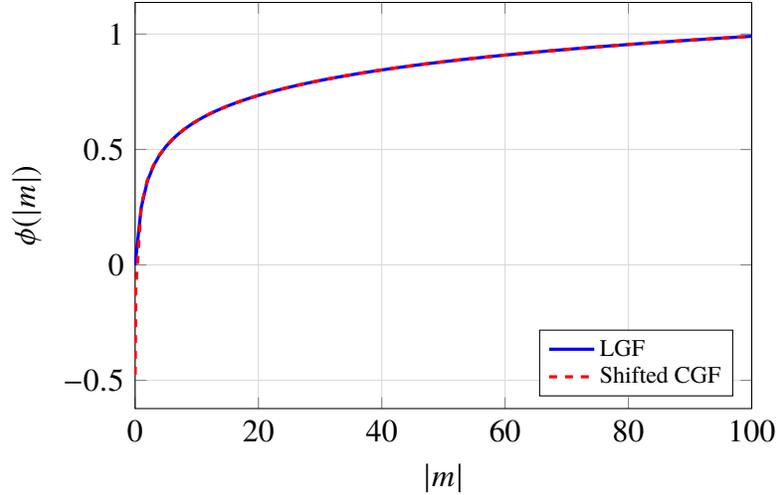
Figure B.1: Comparison of the lattice Green's function to its continuous counterpart for $h = 1$.

where $|\mathbf{m}| = \sqrt{m_1^2 + m_2^2}$ and $\gamma \approx 0.577206$ is the Euler constant. For $|\mathbf{m}| > 30$, dropping the $O(|\mathbf{m}|^{-6})$ renders the truncated expression accurate to at least 10^{-12}.

Defining $(x, y) = h\mathbf{m}$, where h is the grid spacing, one can note that Eq. B.6 implies

$$\phi(\mathbf{m}) + \frac{\log h}{2\pi} \xrightarrow{|\mathbf{m}|\to\infty} \frac{1}{2\pi}\left(\log(\sqrt{x^2 + y^2}) + \gamma + \frac{\log 8}{2}\right) , \tag{B.7}$$

where $1/(2\pi)\log(\sqrt{x^2 + y^2})$ can be recognized as the Green's function for the continuous counterpart of the Poisson problem in 2D. Note that the offset between the discrete and continuous potential is made of two parts: a constant part and a grid dependent part. Also, the behavior near the source differs significantly due to the fact that the continuous potential is $-\infty$ while the discrete one is set to zero.

In fact, it can be shown that the continuous and discrete solutions relate to each other according to

$$u(x(\mathbf{m}), y(\mathbf{m})) \xleftarrow{|\mathbf{m}|\to\infty} \phi(\mathbf{m}) * f(\mathbf{m}) + \frac{1}{2\pi}\left(\log h - \gamma - \frac{\log 8}{2}\right) \sum_{\mathbf{m}\in\mathbb{Z}^2} f(\mathbf{m}) \tag{B.8}$$

and the offset between the two solutions depends not only on the grid spacing but also on the net value of the source.

Consider the potential generated by two collinear charges (magnitudes 2 and -1). Figure B.2 highlights the offset-corrected discrete potential along the centerline and compares it to its continuous counterpart. Note that most of the error is restricted to the vicinities of the charges.
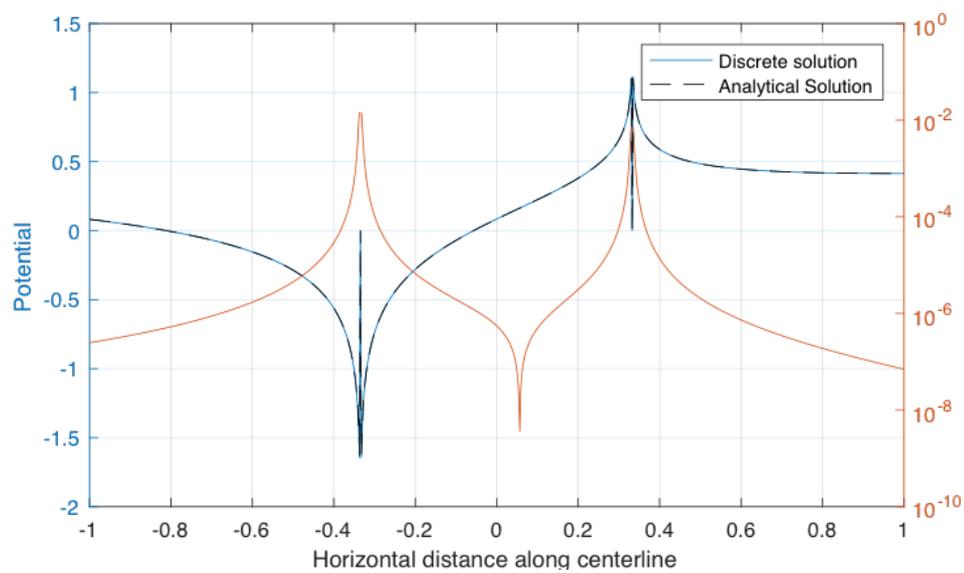
Figure B.2: LGF-based solution corrected to account for the offset. Left axis compares the solution to the exact continuous potential, and right axis shows the local error ($h = 0.001$).

A p p e n d i x   C

DISCRETIZATION OF THE NONLINEAR TERM $U \times \omega$

The nonlinear advection term in Eq. 2.7 correspond to the cross product $\mathbf{u}_r \times \omega'$. The challenge here is the fact that cross product use inputs from the faces of the dual cell. In the present implementation of the code (which follows Colonius and Taira [79]), the following interpolations schemes were used:

- **x-velocity:** primal-cell x faces to dual-cell x faces (primal-cell y faces)

$$u^y_{i,j} = \frac{1}{4} \left(u_{i,j} + u_{i+1,j} + u_{i,j-1} + u_{i+1,j-1} \right) \tag{C.1}$$

- **y-velocity:** primal-cell y faces to dual-cell x faces (primal-cell x faces)

$$v^x_{i,j} = \frac{1}{4} \left(v_{i,j} + v_{i,j+1} + v_{i-1,j} + v_{i-1,j+1} \right) \tag{C.2}$$

- **vorticity:** primal-cell edges to dual-cell x and y faces (primal-cell y and x faces, respectively)

$$\omega^x_{i,j} = \frac{1}{2} \left(\omega_{i,j} + \omega_{i,j+1} \right) \tag{C.3}$$

$$\omega^y_{i,j} = \frac{1}{2} \left(\omega_{i,j} + \omega_{i+1,j} \right) \tag{C.4}$$

Thus, the result of the cross product is given by

$$(\mathbf{u} \times \omega)^x_{i,j} = v^x_{i,j} \omega^x_{i,j} \tag{C.5}$$

$$(\mathbf{u} \times \omega)^y_{i,j} = -u^y_{i,j} \omega^y_{i,j} \tag{C.6}$$

A p p e n d i x   D

THE ERROR INTRODUCED BY THE NONLINEARITIES IN THE ANALYSIS STEP

Under the action of nonlinear forecast and observation functions, a Gaussian initial condition may develop nonzero higher moments in such a way that the state PDF is no longer characterized by just its mean and covariance matrix. The Extended Kalman Filter (EKF) is a first-order approximation to the nonlinear filtering problem, and is the most widely used solution in the industry.

Several second-order extensions to the KF, each of them corresponding to a different choice on how to approximate the presence and action of higher-order moments, have sice been proposed. A popular choice is to suppose that the PDF remains symmetric (third- and higher-order odd central moments are negligible) and concentrated near its mean (even central moment sequence can be truncated at some point). The Gaussian second-order filter was developed by Jazwinski [36], and independently by others researchers. It still assumes the prior to be Gaussian, but retains even central moments up to the fourth-order when solving for the posterior mean. A brief discussion of this algorithm can be found in section 1.4.1.

The goal in this appendix is to compare this second-order filter to its first-order counterpart (EKF), as to better understand the behavior of the error being introduced. Comparing the respective analysis equations written in the representers formulation (inversion in the measurement space), one can see that the second-order representer magnitude b_k^* can be related to EKF counterpart b_k by

$$\left[I + \frac{1}{2}\left(R + \left[\frac{\partial h}{\partial x}(\hat{x}_k)\right]\hat{C}_k\left[\frac{\partial h}{\partial x}(\hat{x}_k)\right]^T\right)^{-1}\left[\partial^2 h\hat{C}_k^2\partial^2 h\right]\right]b_k^*$$

$$= b_k - \frac{1}{2}\left(R + \left[\frac{\partial h}{\partial x}(\hat{x}_k)\right]\hat{C}_k\left[\frac{\partial h}{\partial x}(\hat{x}_k)\right]^T\right)^{-1}\left[\partial^2 h\hat{C}_k\right] \tag{D.1}$$

$$b_k^* = \left[I - \frac{1}{2}\left(R + \left[\frac{\partial h}{\partial x}(\hat{x}_k)\right]\hat{C}_k\left[\frac{\partial h}{\partial x}(\hat{x}_k)\right]^T\right)^{-1}\left[\partial^2 h\hat{C}_k^2\partial^2 h\right] + O(\hat{C}_k^2)\right]^{-1} \times$$

$$\left(b_k - \frac{1}{2}\left(R + \left[\frac{\partial h}{\partial x}(\hat{x}_k)\right]\hat{C}_k\left[\frac{\partial h}{\partial x}(\hat{x}_k)\right]^T\right)^{-1}\left[\partial^2 h\hat{C}_k\right]\right). \tag{D.2}$$

If we assume

$$\left\| \frac{1}{2} \left(R + \left[\frac{\partial h}{\partial x}(\hat{x}_k) \right] \hat{C}_k \left[\frac{\partial h}{\partial x}(\hat{x}_k) \right]^T \right)^{-1} \left[\partial^2 h \hat{C}_k^2 \partial^2 h \right] \right\| < 1 \qquad \text{(D.3)}$$

then

$$b_k^* = b_k - \frac{1}{2} \left(R + \left[\frac{\partial h}{\partial x}(\hat{x}_k) \right] \hat{C}_k \left[\frac{\partial h}{\partial x}(\hat{x}_k) \right]^T \right)^{-1} \left[\partial^2 h \hat{C}_k \right]$$

$$+ \frac{1}{2} \left(R + \left[\frac{\partial h}{\partial x}(\hat{x}_k) \right] \hat{C}_k \left[\frac{\partial h}{\partial x}(\hat{x}_k) \right]^T \right)^{-1} \left[\partial^2 h \hat{C}_k^2 \partial^2 h \right] \times$$

$$\left(b_k - \frac{1}{2} \left(R + \left[\frac{\partial h}{\partial x}(\hat{x}_k) \right] \hat{C}_k \left[\frac{\partial h}{\partial x}(\hat{x}_k) \right]^T \right)^{-1} \left[\partial^2 h \hat{C}_k \right] \right)$$

$$+ O(\hat{C}_k^2). \qquad \text{(D.4)}$$

The first line in Eq. D.4 represents the $O(1)$ term, while the second line represents the $O(\hat{C}_k)$ term. Note that the error in the $O(1)$ term is indistinguishable from a bias in the observation function, and it scales with h_{xx} being fairly independent on $\hat{C}_k$. This conclusion seems to be in agreement with the results presented in Fig. 3.10.

Another particularly famous second-order approach to the nonlinear problem is the Unscented Kalman Filter (UKF), which avoids the need for explicitly computing higher-order derivatives of the forecast and observation models by introducing a sampling strategy that is known as the unscented transformation. A brief description of this methodology can be found in section 1.4.1, and further details can be found in Wan and Der Merwe [47] .